修士論文

セグメントルーティングの階層化に関する研究

1710193　三島 航

主指導教員　篠田陽一 教授
審査委員主査　篠田陽一 教授
審査委員　知念賢一 特任准教授
丹康雄 教授
Razvan Beuran 特任准教授

北陸先端科学技術大学院大学
先端科学技術研究科
（情報科学）

平成 31 年 2 月

概 要

インターネットの利用形態が多様化する中で、利用者の要求もまた多様化と複雑化を続けている。セキュリティ対策や負荷分散・従量課金等の要求に応じ、ファイアウォール・ロードバランサ・Deep Packet Inspection（DPI）など、特定の機能をネットワーク上で提供するミドルボックスがネットワークの管理者により運用され、利用者へと機能を提供している。サービスチェイニングは、複数の機能が存在するネットワークにおいて、特定の通信を複数の機能へ通過させることで一連のサービスとして提供する技術である。その実現のため、宛先や送信元だけでなくサービス単位での経路制御が求められている。実現手法の 1 つとして、柔軟な経路制御技術であるセグメントルーティングが提案されている。セグメントルーティングは経路の各転送対象をセグメントという単位で表現し、ソースルーティングにより経路制御を実現する技術である。既存の MPLS による経路制御と比較し、管理コストの削減や移行の容易性から普及が進んでいる。

NFV の普及やクラウドコンピューティングの発展に伴い、多くの AS で固有のネットワーク機能が提供されている。各 AS が固有の機能を持ち、利用者の選択肢が拡大する中で、複数の AS を連携させたサービスチェイニングの実現が期待される。しかし、既存のセグメントルーティングでは AS ごとの独立と複数 AS の連携した運用の両立が不可能である。そのため、AS ごとの経路広告範囲・ポリシー・ID 空間等の設定の独立が行えないことによる管理コストの増加や、ネットワークの拡大に伴う規模追従性・安定性の課題が生じる。

そこで本研究では、AS 連携環境へのセグメントルーティングの適用を目的とし、セグメントルーティングにサブドメインを導入した階層型セグメントルーティングを提案した。階層型セグメントルーティングは、サブドメイン単位でのドメイン分割を行うことで AS ごとの設定独立と規模追従性・安定性の向上を実現する。その上でコントロールプレーンを階層的に構成し、複数のドメインを束ねることで一連の経路計算を行う。この手法により、サブドメインごとの管理構造・設定の分割が可能となるため、異なる AS が連携した環境においても管理の複雑性の削減や規模追従性・安定性の向上が実現できる。

サブドメインを分割した上で一連の経路制御を実現するため、ドメイン全体の経路計算主体となる階層型 SR PCE を考案した。階層型 SR PCE では下位層の SR PCE が各サブドメインを管理し、上位層の SR PCE がドメイン全体の経路計算を行う。

また、階層的セグメントルーティングにおけるドメインを超えた経路制御を実現するため、転送時のサブドメイン越え手法と、経路構築に用いるポリシー設定手法により複数のモデルを考案した。その上で、サブドメイン越えとポリシー設定のモデルを選出し、階層型 SR PCE の設計・実装を行った。実装した階層型 SR

PCE の性能評価として、複数の AS が存在するネットワークにおいて経路構築時間の計測による規模追従性の評価を行なった。また、階層型セグメントルーティング全体の評価として、サブドメイン分割による AS ごとの管理分割の実現と規模追従性・安定性の性能評価を行なった。それぞれの性能評価の結果から、課題であった AS 連携環境における管理コストの増加や、規模追従性・安定性の低下が解決されたことを示した。

階層的セグメントルーティングの提案により、本研究で対象とした AS 連携環境におけるセグメントルーティングの課題を解決した。階層型セグメントルーティングを適用することで、複数の AS 連携によるサービスチェイニングの提供を始め、より大規模かつ柔軟な経路制御が実現できる。

目 次

図 目 次

表 目 次

第1章　はじめに

本章では、本研究の背景・目的と本論文の構成を述べる。

1.1　背景

インターネットの利用形態が多様化する中で、利用者の要求もまた多様化と複雑化を続けている。セキュリティ対策や負荷分散等をネットワーク上の機能として利用するため、Deep Packet Inspection（DPI）・ファイアウォール・ロードバランサ等のミドルボックスが運用されており、要求に応じて利用者へと提供される。サービスチェイニング [1] は、これらの機能が存在するネットワークにおいて、特定の通信を必要な機能へ通過させることで一連のネットワークサービスとして提供する技術である。ネットワーク機能を汎用サーバ上で実現する Network Function Virtualization（NFV）[2] の提案により、ネットワーク機能の提供がより安価かつ容易に実現可能となったことを受け、サービスチェイニングの重要性は今後も更に高まることが予想される。

サービスチェイニングを提供するためには、宛先や送信元単位ではなくサービス単位での柔軟な経路制御が求められる。その実現手法の 1 つとして、セグメントルーティング [3] が提案されている。セグメントルーティングは既存の Multi-Protocol Label Switching（MPLS）におけるラベル配布を用いた経路制御よりも簡素なプロトコル構成による管理コストの削減や、MPLS からの移行の容易性などの利点を持つ。そのため、Autonomous System（AS）内部におけるサービスチェイニングや、通信・アプリケーション単位でのトラフィックエンジニアリング [4] などでの利用が期待されている。

クラウドコンピューティングの発展によるパブリッククラウドの増加や ISP が提供するサービスの多様化に伴い、様々な AS でネットワークサービスが提供されるようになった。複数の AS が存在するネットワーク環境において、特定の AS に固有な機能同士をサービスチェインとして繋ぎ提供することで、利用者に新たな価値を提供できる。しかし、ISP 間の連携やマルチクラウド連携によるサービスチェイニング等、複数の AS が存在するネットワークへの適用を考えた際、既存のセグメントルーティングには管理の複雑性や規模追従性・安定性の課題が生じる。

1.2 目的

複数の AS が存在する環境では、AS ごとの管理構造の独立や、各 AS が個別の IGP 設定・ポリシー・ID 空間を持つことが求められる。しかし複数の AS をセグメントルーティングにより連携させる場合、既存手法では同一のドメインに含めて運用する必要があるため、AS ごとの分離が行えず管理の複雑化や運用コストの増加が生じる。また、セグメントルーティングでは同一のドメインに属する各ノードが全セグメントの識別子情報を保持する必要があるため、ネットワークの規模拡大に伴い経路情報の増加とセグメント情報伝搬範囲の増加に伴うネットワーク安定性の低下が生じる。

本研究では複数の AS が存在するネットワークにセグメントルーティングを適用するため、セグメントルーティングの管理構造の分割による複雑性の解消と、規模追従性・安定性の向上を実現する。複数 AS が連携した環境へとセグメントルーティングを適用することで、マルチクラウド連携や ISP を越えたサービスチェイニングの提供が可能となる。

1.3 本論文の構成

本論文は、本章を含めて 7 章から構成される。2 章では、柔軟な経路制御の実現手法であるセグメントルーティングの概要とその基本動作、および AS 連携における課題について述べる。3 章では、セグメントルーティングの拡張や規模追従性に関連する研究、またルーティングプロトコルにおける大規模化に関する研究について説明し、2 章で述べた課題に対して不足している機能を明らかにする。4 章では、2 章で述べた課題を解決するためセグメントルーティングに階層構造を導入した階層型セグメントルーティングを提案し、実現に必要となる機能を議論すると共に、2 つの軸で実現方式のモデル化を行う。5 章では、提案手法に基づくシステムの設計と実装について述べる。6 章では、提案手法の評価実験・性能評価・既存手法との比較について述べる。7 章では、本論文のまとめと今後の課題について述べる。

第2章　柔軟な経路制御の実現手法と課題

本章では、ネットワークにおける柔軟な経路制御の重要性とその実現手法であるセグメントルーティングの概要、適用先の拡大とそれに伴い生じる課題について述べる。

2.1　経路制御の要求とトラフィックエンジニアリング

トラフィックの種別や回線状況、遅延に応じたリソース割り当てなどの経路操作によるネットワーク最適化技術をトラフィックエンジニアリングと呼ぶ。トラフィックエンジニアリングの技術は、輻輳回避の他に、遅延やジッタを考慮した通信品質の保証、ロードバランシング、トランジットを考慮した経由 AS の制御などにも用いられている。

インターネットの利用形態が多様化し、利用者の要求が複雑になる中で、経路制御技術への要求も複雑化を続けている。その代表的な例としてサービスチェイニングがあげられる。サービスチェイニングは、特定の通信をネットワーク上に存在するファイアウォールやロードバランサなどの必要な機能群へと通過させることで、一連のサービスとして提供する技術である。ここで提供される一連のサービスをサービスチェインと呼ぶ。契約等に基づいて特定の送信元・種類の通信に対しサービスを提供するため、従来の宛先 IP アドレス単位での経路制御ではなく、送信元アドレスや提供されるコンテンツ・サービス単位での経路制御が必要となる。このように、複雑化した要求や利用形態に応じるため、宛先単位での経路制御ではなくより柔軟かつ細かい粒度での経路制御が求められている。

本節では、インターネットにおける経路制御技術を AS 内・AS 間に分類し、その特徴と課題を述べる。

2.1.1　AS 内における経路制御技術

AS 内部におけるトラフィックエンジニアリングの代表的な実現手法として、データリンク層プロトコルによる実現手法と、MPLS によるカプセリングを利用する実現手法が存在する。

データリンク層プロトコルによるトラフィックエンジニアリングとしては、フレームリレーや Asynchronous Transfer Mode（ATM）による実現が挙げられる。フレームリレーや ATM においては、負荷やポリシーにしたがって仮想回線を切り替えることにより、トラフィックエンジニアリングを実現できる。また、輻輳制御技術としては、フレームリレーでは Forward Explicit Congestion Notification（FECN）と Backward Explicit Congestion Notification（BECN）による輻輳状態通知の仕組みが、ATM では Available Bit Rate（ABR）による輻輳状態通知の仕組みがそれぞれ存在する。フレームリレーや ATM による実現では、宛先 IP アドレスに捉われない柔軟な経路制御が実現できるが、各プロトコルに応じた専用機材を用いネットワークを作成する必要がある点や、ヘッダサイズによる帯域消費が生じる点、複数プロトコルの同時運用により、運用者の学習コストや運用障害時の問題切り分けコストが高くなる点といった課題がある。

一方、専用のデータリンク層プロトコルを使用せず、MPLS によるカプセリングを用いる方式として MPLS-TE が存在する。MPLS-TE には、ラベル配布・経路生成に Label Distribution Protocol（LDP）を用いる Constraint-based Routing LDP（CR-LDP）方式と Resource reSerVation Protocol を用いる REVP-TE 方式があるが、機能がほぼ同一であるため CR-LDP は廃止されている。RSVP-TE は帯域予約プロトコルである RSVP を拡張し、ラベル配布や経路設定機能を持たせることで、トラフィックエンジニアリングを行う方式である。この方式ではラベル配布やトンネル作成のためのシグナリング専用プロトコルである RSVP と、ネットワークの帯域情報やリンクのメトリックなどの情報を伝送する Open Shortest Path First Traffic Engineering（OSPF-TE）や Intermediate System to Intermediate System Traffic Engineering（ISIS-TE）などのルーティングプロトコルが分離しており、管理が煩雑という課題がある。

2.1.2 AS 間における経路制御技術

AS 間のトラフィックエンジニアリング実現手法としては、Border Gateway Protocol（BGP）を用いた手法がある。BGP にはパス属性と呼ばれる経路を制御するための属性が存在する。隣接する AS へ対して複数に分割した経路情報へ Multi Exit Discriminator（MED）などの異なるパス属性を持たせて広報することで、宛先 IP アドレス毎の経路制御を実現できる。BGP を用いた手法では、経路制御が IP アドレスや AS ごとの単位となり詳細な経路制御が行えないというデメリットがある。AS 間でのより細かい経路制御を実現するために、MPLS を用いた VPN 技術である MPLS-VPN が存在する。MPLS-VPN では、各ルータは Virtual Routing and Forwarding（VRF）により VPN のルーティングテーブルを分割し管理を行う。そして各 VRF の情報に Route Distinguisher（RD）・Route Target（RT）と呼ばれる識別情報をつけ、BGP4 の拡張技術である Multi Protocol BGP（MP-BGP）により RT をピアに広報することで、ルーティング情報を共有し、AS を越えた MPLS

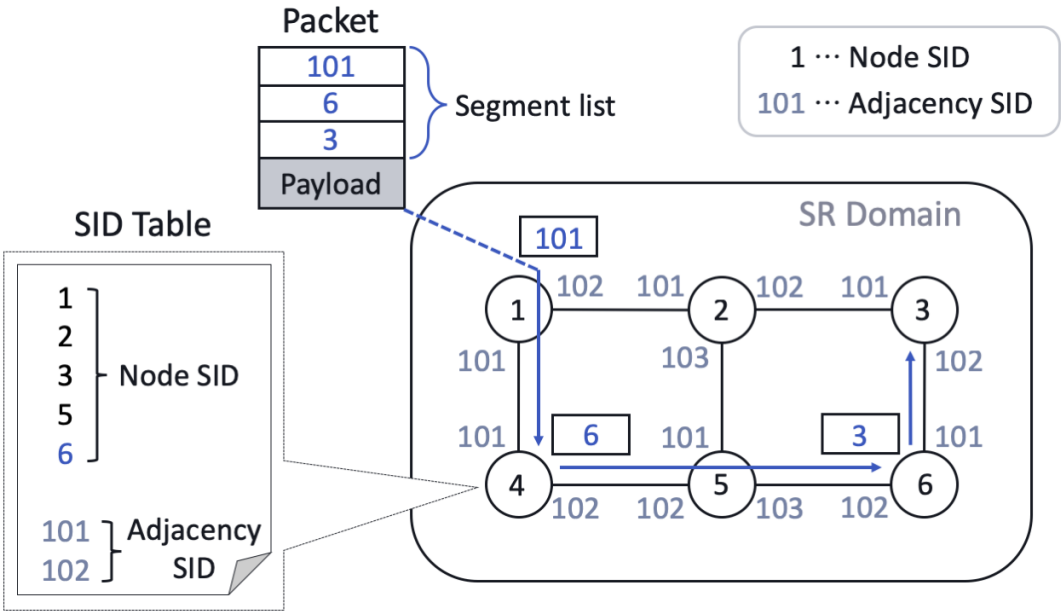

図 2.1: セグメントルーティングの概念

通信を実現する。MPLS-VPN を用いた手法では、経路制御を行うために MPLS 対応ルータが必要となるという課題がある。また、BGP を用いた手法・MPLS-VPN を用いた手法の双方において、経路制御のために AS 内部のプロトコルと AS 間の 2 種以上のプロトコルのが必要であり、学習コストの増大や、障害時の原因切り分けが複雑になるという課題が生じる。

2.2　セグメントルーティングの概要

第 2.1 節で述べた課題を解決するため、セグメントルーティングが考案されている。セグメントルーティングとは、転送に用いるネットワークのノードや隣接関係などの各要素をセグメントと呼ばれる単位で扱い、経由先や宛先として指定することで柔軟な経路制御を実現する技術である。2013 年 11 月に IETF のワーキンググループである Source Packet Routing in Networking（SPRING）[5] により提案され、仕様策定が進められている。セグメントルーティングはコントロールプレーンとデータプレーンが分離しており、経路決定を行うコントロールプレーンのプロトコルとして OSPF・IS-IS が、パケットの転送を行うデータプレーンのプロトコルとして MPLS や IPv6 の拡張ヘッダの利用が提案されている。データプレーンに MPLS を用いるものを SR-MPLS、IPv6 を用いるものを SRv6 と呼称する。

セグメントルーティングの概念図を図 2.1 に示す。セグメントルーティングに参加するノードの集合を SR ドメインと呼ぶ。セグメントルーティングでは各セグメントを Segment ID（SID）と呼ばれる識別子で識別する。SR ドメインでの送信元

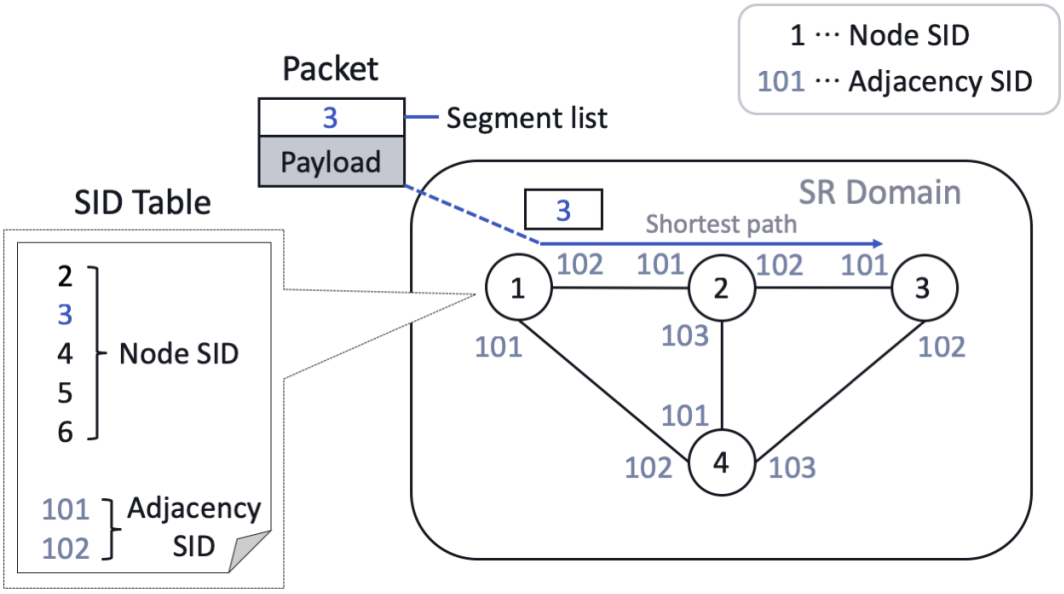

図 2.2: Node Segment

となるノードが経路を指定するソースルーティングを採用している。経由する全セグメントの SID リストであるセグメントリストをパケットに付加することで経路を指定する。各ノードは転送の対象となる全ての SID 情報を SID テーブルとして保持しており、パケットに付加されたセグメントリストの先頭要素に従い転送先を決定する。図 2.1 では、パケットに<101, 6, 3>の順でセグメントの指定されたセグメントリストが付加されている。セグメントルーティングで扱われる代表的なセグメントには、ノードを表現する Node Segment とノード間の特定の隣接関係を表現する Adjacency Segment が、また拡張性の向上のため Binding Segment が存在する。

2.2.1 Node Segment

Node Segment は特定のノードを表すセグメントであり、Prefix Segment とも呼ばれる。Node Segment の概要を図 2.2 に示す。

以降の文章では Node SID 1 のノードをノード 1、Node SID 2 のノードをノード 2 というように番号で指定を行う。Node Segment の SID が指定された場合、その Node SID が示すノードへ、アンダーレイネットワークにおける最短経路に従いパケット送出が行われる。図 2.2 の例では、Node SID 3 が示すノード 3 へ向け、アンダーレイネットワークでの IGP における最短経路に従い転送が行われている様子を示している。Node SID は SR ドメイン内でグローバルユニークな値であり、SR ドメインに属するあらゆるノードが全 Node SID 情報を保持する必要がある。

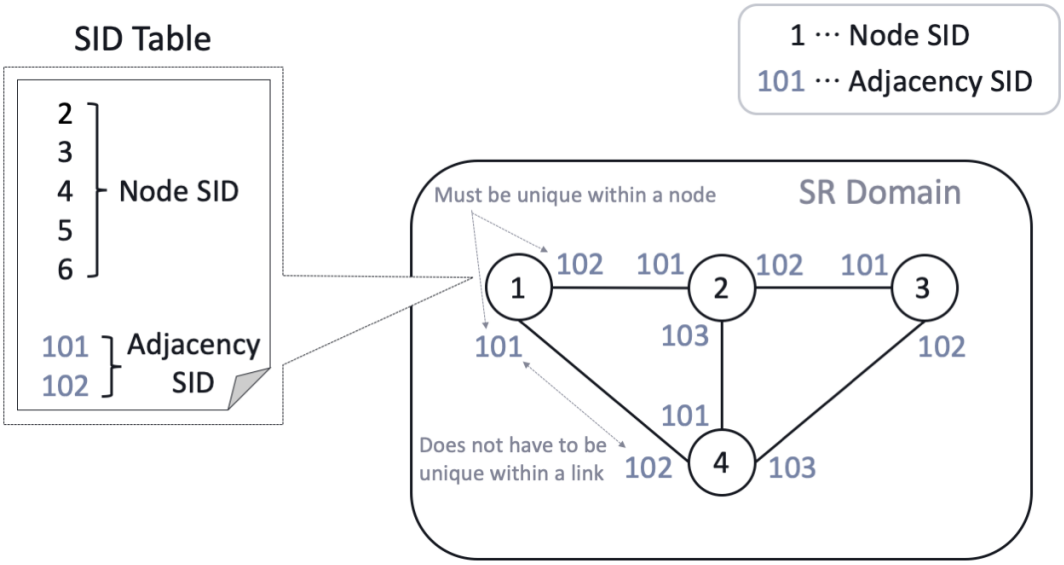

図 2.3: Adjacency Segment

2.2.2 Adjacency Segment

Adjacency Segment は、任意のノード内部における隣接関係を表すセグメントである。Adjacency Segment の概要を図 2.3 に示す。

Adjacency SID が指定された場合、そのノードにおいて Adjacency SID が示すインターフェースからパケットの送出が行われる。Adjacency SID は各ノードにおいてインターフェースを識別するためのローカルユニークな値である。そのため図 2.3 の通り、Adjacency SID は各ノード内で一意であるが、異なるノードでは一意性は保証されておらず同じ値が用いられる可能性がある。また、同じ隣接関係を表す Adjacency SID であっても、対向のノードとの一意性はなく、必ずしも同一の値とはならない。Adjacency Segment によるパケット転送の様子を図 2.4 に示す。

図 2.4 は、ノード 1 からノード 3 へとパケットを転送する様子を示している。その際、Adjacency SID を利用しノード 2 とノード 4 間のリンクを経由させる。この場合のセグメントリストは<102, 103, 103>となる。パケットを受け取った際、ノード 1 はセグメントリストの先頭要素が 102 であるため Adjacency SID 102 の示すインターフェースを送出先として決定する。そしてセグメントリストの先頭要素を次のセグメントである 103 へ変更し送出を行う。ノード 2 でもセグメントリストの先頭要素となった 103 を確認し、Adjacency SID 103 の示すインターフェースから送出を行う。ノード 4 も同様に先頭要素を確認し送出を行う。Adjacency SID は各ノードにおいてインターフェースを識別するためのローカルユニークな値で

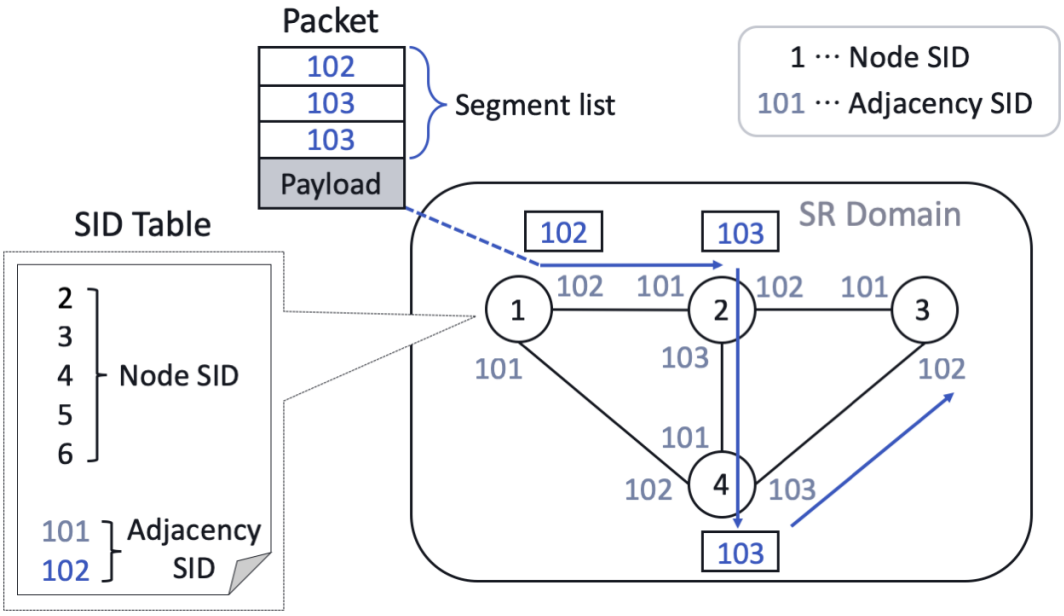

図 2.4: Adjacency Segment の動作

あるため、図 2.4 のように、 1 つの SR ドメインの中において複数のノードで同一の Adjacency SID が利用される可能性がある。

2.2.3 Binding Segment

Binding Segment[3] は、セグメントルーティングの拡張性を高めるため提案されたセグメントである。ノードを識別する Node Segment や隣接関係を識別する Adjacency Segment と異なり、任意のポリシーを識別する。Binding Segment は SR ドメインで一意な SID としても、ノード内のみで一意な SID としても利用可能であり、その割り当ても動的・静的の双方が提案されている。Binding Segment 利用方法の 1 つとして、そのポリシーを満たすよう運用される RSVP-TE・IP/UDP・GRE 等のトンネルや、光回線等の任意のインターフェースへと結合することが提案されている。

Binding Segment の利点として、任意のポリシーと結びつけることが可能なため拡張性が高い点や、ポリシーそのものと結びつけるという性質から、経路の変更等の具体的な経路状態が隠蔽できる。また、トンネルや経路に結びつける場合は具体的な経路状態の隠蔽が可能となる。

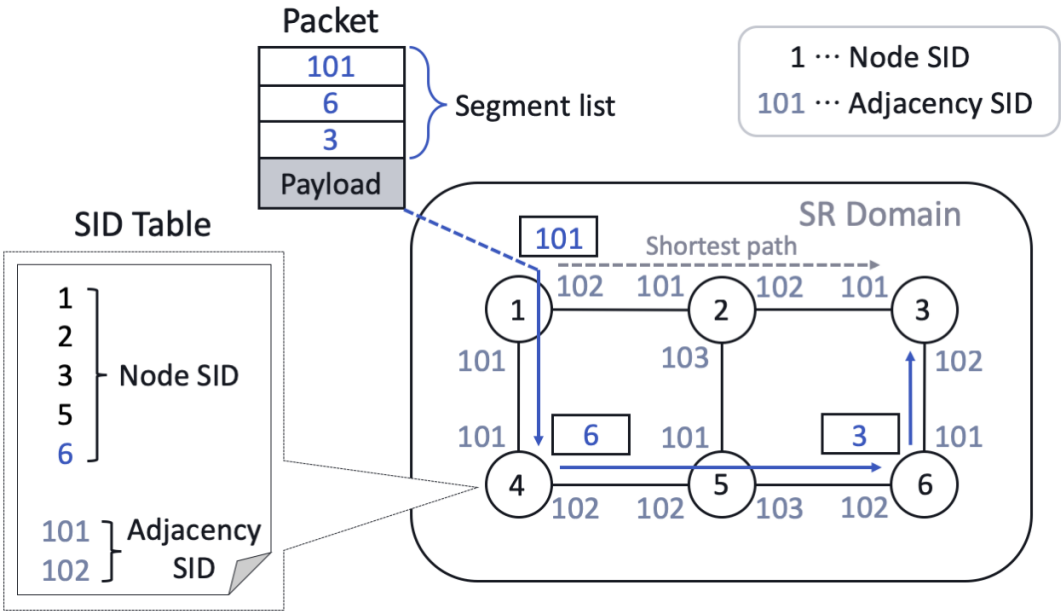

図 2.5: セグメントルーティングによるパケット転送

2.2.4 セグメントルーティングによる転送例

パケットを受信したノードは、セグメントリストの先頭が Node SID である場合は経路表における宛先への送出インターフェースへ、Adjacency SID である場合には該当隣接関係が存在するインターフェースへとパケットを転送する。また、セグメントリストの先頭が自らを示す Node SID や、Adjacency SID であった場合には先頭要素を次のセグメントに変更する。そのセグメントが最終要素であった場合、セグメントルーティングによる転送が終了する。

図 2.5 は SR ドメインに属するノード 1 が<101, 6, 3>というセグメントリストの付加されたパケットを受信した例を示している。ノード 1 からノード 3 への IGP における最短経路は、ノード 2 を経由する経路である。この場合、セグメントルーティングでは以下の手順でパケット転送が行われる。(1) パケットを受信した際、ノード 1 はセグメントリストの先頭要素を確認する。先頭要素は 101 という Adjacency SID であるため、ノード 1 はセグメントリストの次の要素である 6 を先頭とし、101 が示すインターフェースからパケットを送出する。(2) パケットを受け取ったノード 4 は、セグメントリストの先頭要素である SID 6 が示すノード 6 への IGP における最短経路を利用しパケットを送出する。ノード 5 は先頭要素が隣接するノード 6 であるため、先頭要素を 1 つ先へ進めた後にノード 6 へ送信を行う。(3) ノード 6 はセグメントリストの先頭要素である 3 を確認し、Node SID 3 が示すノード 6 への最短経路を利用しパケットを送出する。その後ノード 3 がパケットを受信することでセグメントルーティングによる転送が終了する。

2.2.5　セグメントルーティングの利点

第 2.2.4 節でセグメントルーティングにより柔軟な経路制御が実現可能であることを示した。RSVP-TE 等の他の経路制御技術と比較し、セグメントルーティングの利点として、LDP や RSVP による Label Switched Path（LSP）・Forwarding Equivalence Class（FEC）単位でのトンネリング状態の保持が不要であり、Node と Adjacency のみの状態保持に削減できること、また、専用のシグナリングプロトコルが必要である RSVP-TE と比べ、コントロールプレーンに IGP のみを用いたシンプルな構成が可能である点から、管理コストの軽減が可能となる。

2.2.6　SR PCE

セグメントルーティングでは Software Defined Network（SDN）[6] への対応等を目的とし、Path Computation Elements（PCE）[7] による情報の管理である SR PCE[8] が提案されている。本論文では、SR ドメインに対するポリシー設定対象の明確化と設定手法の簡素化のため、SR PCE の利用を想定する。

PCE とは、主に MPLS においてネットワーク全体の経路計算を特定のノードで行う技術である。PCE には、発行した LSP の状態保持・同期による最適化を行う Stateful PCE と状態保持・同期を行わない Stateless PCE が存在する [9]。PCE を用いないネットワーク運用では各ルータがネットワーク全体の状態を保持し分散的に経路計算を行うが、この手法では各ルータの計算能力やドメイン分割の制約により、トラフィック量や遅延の全体最適を実現するようなポリシー設定やドメインを越えた経路などの複雑な経路計算が不可能という問題が生じる。これに対し、PCE を用いた管理により、ネットワーク全体のトポロジや経路の状態を考慮した全体最適を実現するポリシー設定が可能となる。

SR PCE は PCE をセグメントリストに適用し、経路計算のソフトウェア制御や全体最適の実現などの利点を SR に導入するための技術である。セグメントリストとなる SR-TE LSP の計算を行うための計算主体として、SPRING より提案され、仕様策定が進められている。SR PCE の概念を図 2.6、図 2.7 に示す。

PCE の処理は、大きくリンクステート収集と経路確立に分けられる。PCE に経路計算を依頼するノードを Path Computation Client（PCC）と呼ぶ。図 2.6 に PCC から PCE へのリンクステート配布処理を示す。リンクステート配布処理では各 PCC が PCE とのセッションを確立し、トポロジ情報と予約可能帯域やメトリック等のトラフィックエンジニアリング情報が含まれたリンクステート情報を送信する。PCE にはリンクステート情報を Traffic Engineering Database（TED）と呼ばれるデータベースへと格納を行い、経路情報の共有が行われる。

図 2.7 にセグメントリスト確立処理を示す。まず経路確立を希望する PCC が PCE へと経路確立要求である Path Computation Request（PCReq）を送信する。PCReq を受信した PCE は TED のリンクステート情報とあらかじめ設定されたポ

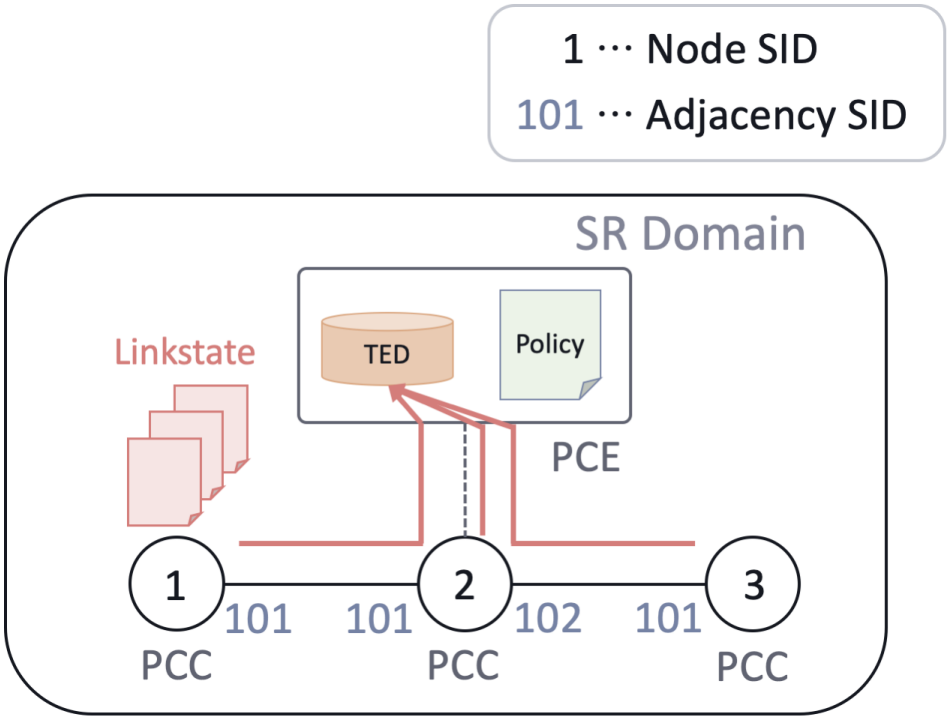

図 2.6: SR PCE によるリンクステート収集

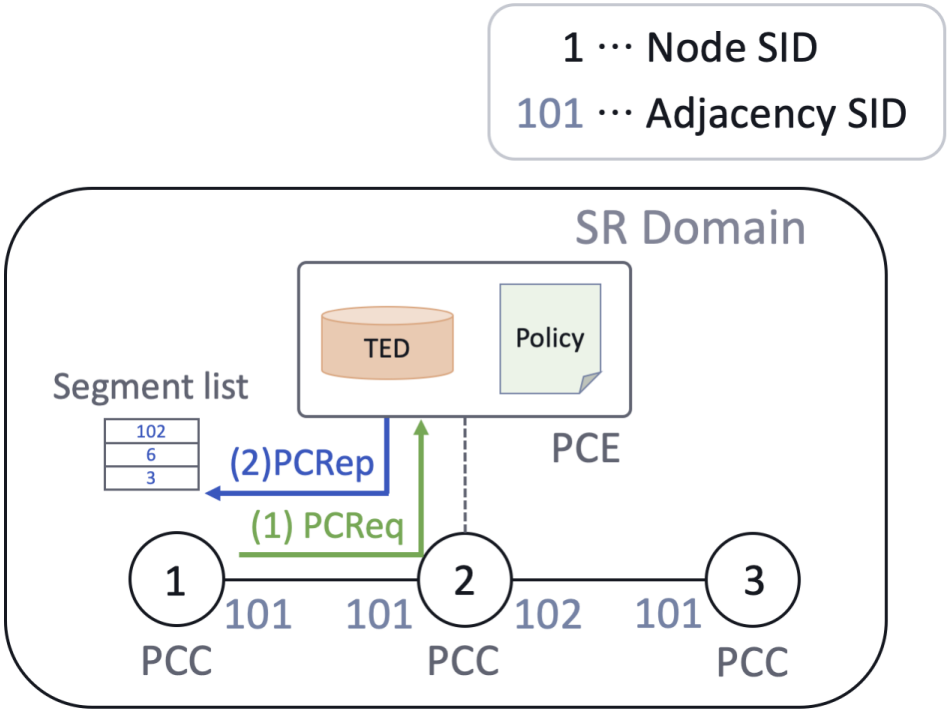

図 2.7: SR PCE によるセグメントリスト確立

リシーに基づき、Constrained Shortest Path First（CSPF）[10] 計算により制約付き経路を生成する。その後、Path Computation Reply（PCRep）として、PCC へ経路情報の配布を行う。SR PCE であれば経路情報としてセグメントリストを配布する。

このように、SR PCE を利用することでネットワーク全体の経路計算が可能となり、複雑なポリシー設定やネットワーク全体の一元管理が実現できる。既存の PCE では、図 2.6 で示したリンクステート共有に Border Gateway Protocol Link-State（BGP-LS）[11] が、図 2.7 で示した経路確立処理に Path Computation Element Protocol（PCEP）[12] が用いられる。

2.3 セグメントルーティング適用先の期待とその課題

第 2.2.5 項で述べた利点から、セグメントルーティングは柔軟な経路制御と簡易な運用を実現する技術としてその適用が期待されている。しかし、大規模利用や複数 AS の連携する環境など一部の環境においては課題が生じる。本節ではセグメントルーティングの適用先の期待と生じる課題について述べる。

2.3.1 AS を越えた End-to-End のサービスチェイニング

第 2.1 項で述べたように、セグメントルーティングの代表的な利用例としてサービスチェイニングがある。既存のセグメントルーティングによるサービスチェイニングは、単一の AS 内に存在する Virtual Network Function（VNF）を利用し実施される。ここで新たな要求として、プライベートクラウドとパブリッククラウドのような異なる AS に存在する VNF を連携させたサービスチェイニングを考える。図 2.8 に複数の AS に VNF が分散し存在した例を示す。

図 2.8 の例では 3 つの AS が存在する。1 つ目は図の右下に存在する ISP-S、2 つ目は図の右上に存在する ISP-D、3 つ目は図の左に存在するパブリッククラウドである。ISP-S は顧客である host1 と host2 を収容している。また、VNF としてファイアウォールを提供している。ISP-D は Web サーバを顧客として収容している。パブリッククラウドは、ファイアウォールと DPI の 2 つの VNF を収容している。左側に示した経路が host1 から Web サーバへの最短経路、右側に示した経路が host2 から Web サーバへの最短経路である。このネットワークにおける ISP-S のサービスとして、顧客が任意の VNF を選択しサービスチェイニングを構築する仕組みを提供することを考える。このサービスの実現により、以下の 3 つの利点が生じる。（1）顧客側の利点として、任意の VNF が利用できることによるネットワーク機能の選択肢の増加と利用の簡易化（2）ISP 側の利点として、顧客へのサービス提供による恒常的な収益の獲得（3）パブリッククラウド側の利点として、ISP を通じた VNF 提供による新規利用者獲得・収益の増加

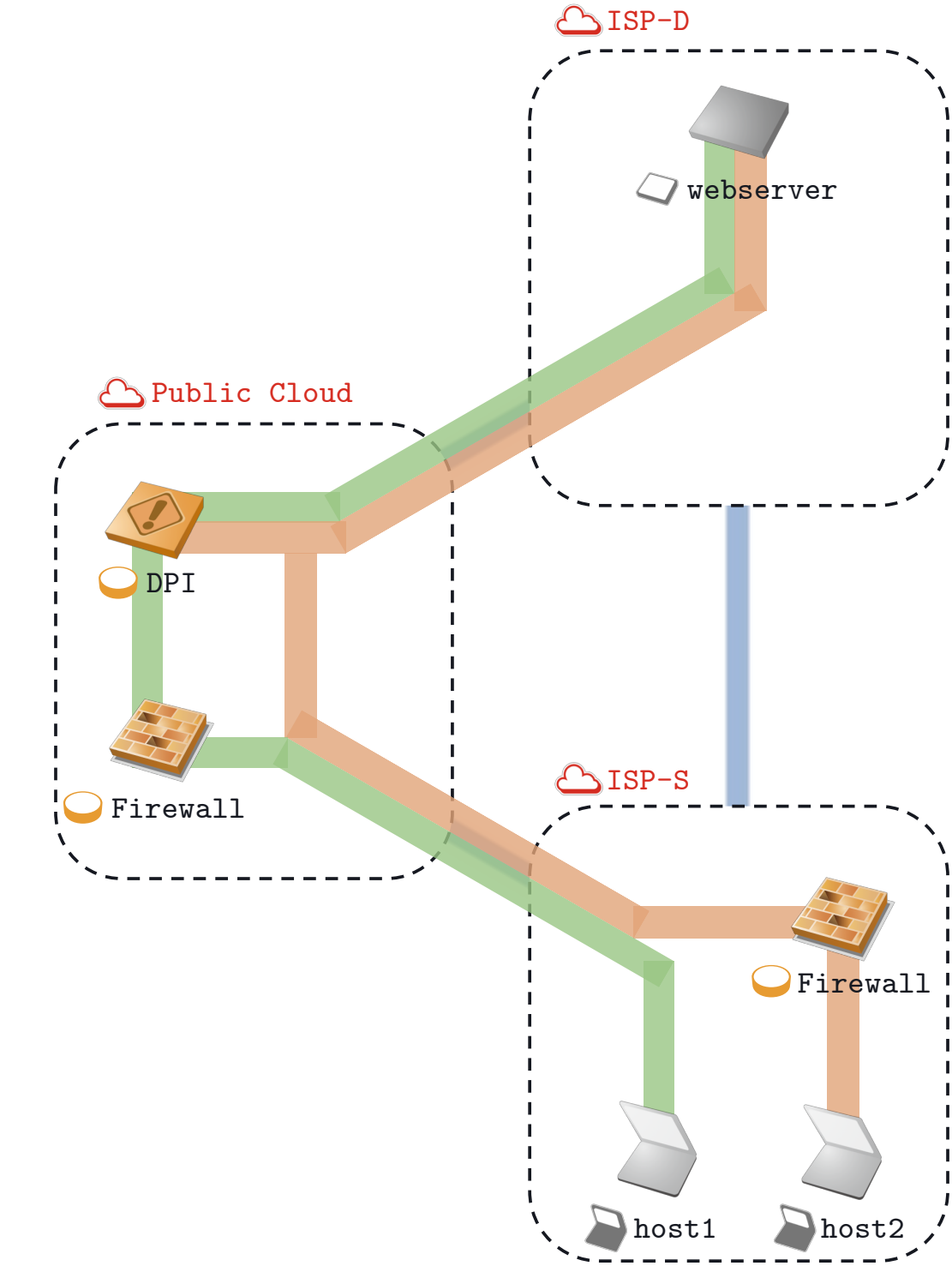

図 2.8: 複数の AS に VNF が分散した環境

具体的な経路制御の例として、図 2.8 の環境で、host1・host2 のそれぞれから Web サーバへの 2 種類のサービスチェインを作成する。(1) host1 は ISP-S の提供するファイアウォールと、パブリッククラウドの提供する DPI を利用する。(2) host2 はパブリッククラウドの提供するファイアウォールと DPI を利用する。図 2.8 では、図 2.8 中の左に示した経路が host1 から Web サーバへのサービスチェイン、右に示した経路が host2 から Web サーバへのサービスチェインである。双方の経路が要件を満たしつつ、AS を越えた制御を実現している。

複数の AS を越えた経路制御を実現することで、例のように AS 単位に捉われない柔軟なサービスの提供が可能となる。AS 連携のように、複数のドメインを統合した一連の経路制御やドメイン単位での抽象化を行うネットワークを Abstraction and Control of Traffic Engineered Networks（ACTN）[13] と呼び、柔軟なサービス提供や経路制御のため、その実現が求められている。しかし、ACTN を満たすような複数の AS を越えた経路制御を実現する場合、既存のセグメントルーティングには次の 2 つの課題が生じる。(1) 複数の AS に一連のサービスチェインを構築するため、異なる AS が同一の SR ドメインに属するよう構築する必要があり、管理の複雑化や運用コストの増大が生じる。(2) セグメントルーティングでは各ノードが SR ドメイン内の全ての Node SID を SID テーブルに保有する必要があるため、複数の AS を 1 つの SR ドメインに収容した場合、経路変更時の SID エントリのアップデートによるネットワークの安定性低下やテーブル量の肥大による規模追従性の課題が生じる。それぞれの課題について、以下の第 2.3.2 節、第 2.3.3 節で述べる。

2.3.2 AS 連携における課題

既存のセグメントルーティングにおいて、AS ごとに設定を独立させたまま同一の SR ドメインに参加させることは不可能である。仮に異なる AS が存在するネットワークにセグメントルーティングを適用するためには、全ての AS を単一の SR ドメインに設定する必要がある。しかしその場合、異なる AS と、IGP 設定や ID 空間をはじめとするネットワーク設定規則の統一が必要となり、設定作業や変更等を隣接する AS と情報交換しながら行う必要があるため、管理の複雑化が生じる。また、各 AS が SR ドメインに存在する全てのノードを他の AS へ広告する必要が生じ、隣接 AS へのノードの隠蔽が行えなくなるという制約が生じる。それに加え、パブリッククラウドが複数の AS とサービスチェイニングを行う場合、全ての AS を 1 つの SR ドメインに包括させる必要があり、直接サービスチェインを構築しない AS 間においてもノード情報や ID 空間を共有する必要性や、予期せぬ AS から通信を受ける可能性が生じるなど、適切な単位での設定分割が行えなくなるという課題が生じる。更に SR-MPLS を利用していた場合、Node SID の一意性の制約から MPLS のラベル空間の共有と、設定規則の共通化が必要となる。

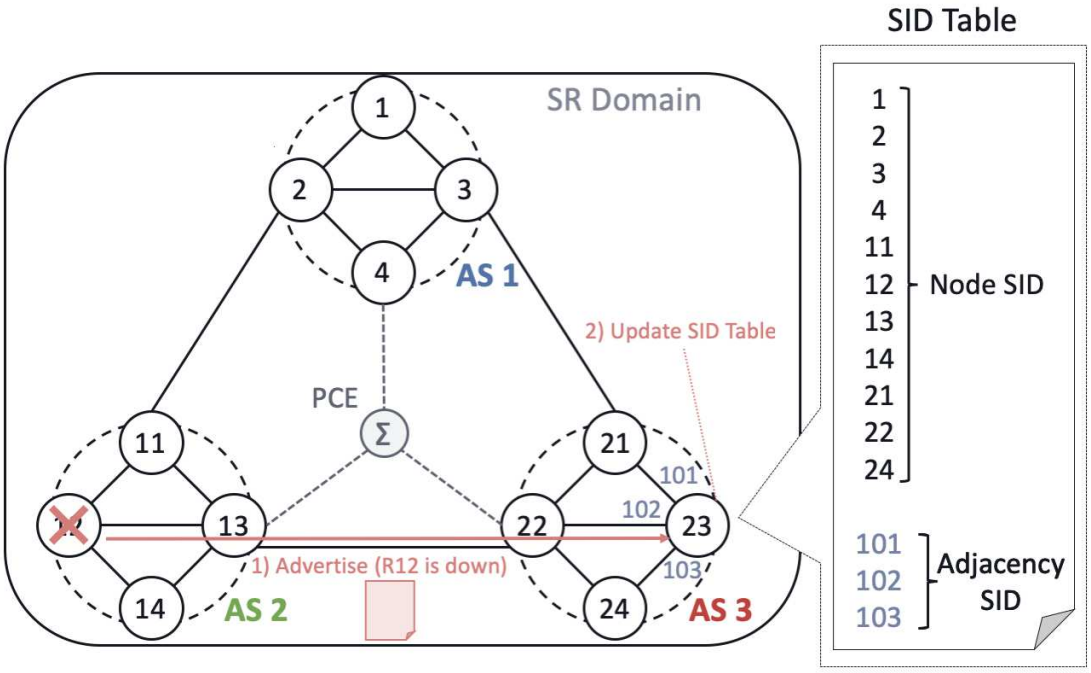

図 2.9: セグメントルーティングの大規模利用に伴う SID テーブルの肥大

2.3.3 大規模利用における課題

第 2.2 節で述べたように、セグメントルーティングはパケットの転送対象としてセグメントを指定することで経路制御を行う技術である。代表的なセグメントとして、Node Segment と Adjacency Segment が存在する。SR ドメインの任意のノードから全てのノードへ通信を行うため、Node Segment は SR ドメイン上で一意な値であり、参加する各ノードが全ノードの Node SID を保有する。そのため大規模なネットワークにおいてセグメントルーティングを適用した場合、SR ドメインに参加するノード数の増加に伴い各ノードの持つべき SID テーブルの情報量が増加することで、メモリ量の消費と SID 情報のアップデートに伴うネットワークの安定性低下により規模追従性に問題が生じる。図 2.9 に、セグメントルーティングの大規模利用に伴い SID テーブルの肥大が生じる様子を示す。

図 2.9 の例では、AS3 に属するノード 23 が同じ SR ドメインに属する AS1 や AS2 のノードの Node SID も保有しており、SID テーブルが肥大化していることが確認できる。また、SID の共有範囲が拡大することでネットワークの安定性の低下が生じる。図 2.9 の例では、AS2 内部のノード 12 に異常が生じトポロジの変更が起きた様子を示している。SID 情報の変更は SR ドメイン全体へ伝搬するため、AS1・AS3 に存在する全てのノードに広告が行われ、SID テーブルの変更が生じる。IP ネットワークでは他 AS のノードはネットワークアドレスにより隠蔽されるため、ノードの追加・削除という単位で経路表の書き換えが起きることは少ない。しか

し、セグメントルーティングでは全てのノードの情報がSIDテーブルに格納されるため、他ASのノード単位でのトポロジの変更により全ノードのSIDテーブルの変更が生じるため、ネットワークの安定性が低下する。

第1章で扱ったように、インターネットにおけるトラフィックの増加や用途の複雑化などの背景から、大規模なネットワークにおいて、複雑な要求を満たすトラフィックエンジニアリングが求められている。その実現手段としてセグメントルーティングを適用するため、複数のASが連携した環境へセグメントルーティングを適用した際に生じる管理の複雑化・コスト増加等による規模追従性低下と規模拡大に伴う安定性低下を、本研究の課題と定める。

第3章　関連研究

本章では、セグメントルーティングにおける拡張研究、また他のルーティングプロトコルにおける大規模化に関する研究について扱い、第 2 章で述べた課題に対して不足する要件を明らかにする。

3.1　セグメントルーティングの拡張研究

セグメントルーティングの適用先の期待に応じ、それぞれの用途に合わせセグメントルーティングを拡張する様々な研究が行われている。本節ではそれらの研究内容について述べる。

3.1.1　Binding Segment による拠点間通信

第 2.2.3 項で扱った Binding Segment の利用先として、セグメントルーティングを用いた拠点間・ドメイン間における通信が提案されている。

図 3.1 に、SPRING において Filsfils らにより提案された、Binding Segment による拠点間通信 [14] を示す。この例では、DC1・Core・DC2 の 3 つのネットワークが存在しており、各ネットワークは個別の SR ドメインとして動作している。この環境において、BSID を用いることで一連の経路制御を行う。Filsfils らはこの手法の利点として、隣接するネットワークで用いられるプロトコルや経路状態の隠蔽により、遅延やジッタ制限等、ポリシー単位での抽象化が可能である点と、BSID の利用によるセグメントリストの削減を挙げている。

この手法から考えられる課題として、異なる AS の連携など、各拠点の管理者が異なる場合のポリシー管理手法の検討が必要となると考えられる。また、新たな経路の作成時、ソースルーティングの利点である送信元のみでの経路生成ではなく、ドメインを越えた複数箇所での経路設定や対応する BSID 生成が必要となることも課題となる。これらの検討から、複数のネットワークを繋いだ経路制御が可能であるという利点を残しつつ、AS 連携に必要となる経路適用・経路発行時の複数箇所での BSID 設定・広告の仕組みや、管理者同士の連携手法の確立が必要となる。

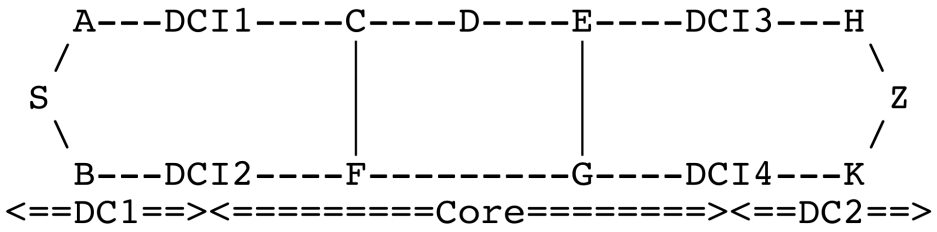

Figure 9: A Simple Datacenter Topology

図 3.1: Binding Segment を用いた拠点間通信（出典:Filsfils, et al., Segment Routing Policy for Traffic Engineering, 2017.）

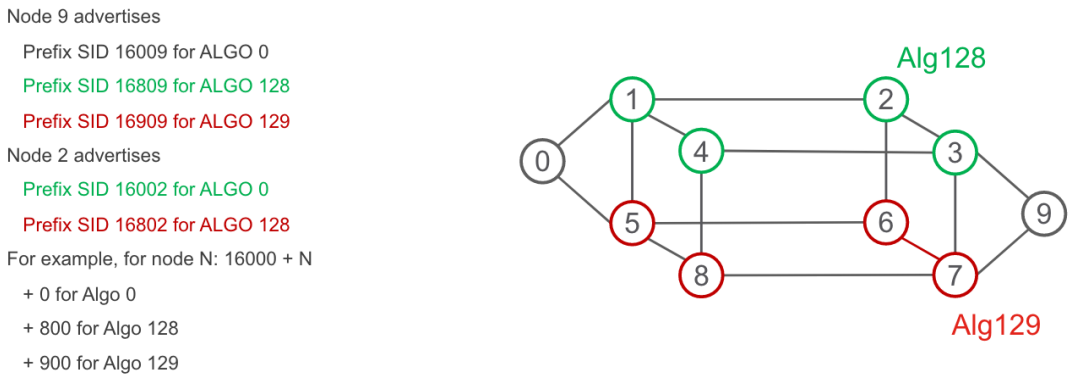

図 3.2: Flex Algorithm の適用例（出典：Filsfils, SR IGP Flex Algo, 2018.）

3.1.2 Flex Algorithm

Flex Algorithm[15]（以下、Flex Algo）は、セグメントルーティングにおいて制約付きパス計算を実現するためのアルゴリズムである。Flex Algo では、通常の Shortest Path First（SPF）計算による最短経路ではなく、遅延や帯域保証、特定のノードやリンクの除外などの制約に従って経路計算を行い、転送を実現する。具体的な手法としては、Flex Algo に参加するノードは通常の Node SID とは異なる Node SID を決定し、Flex Algo を識別する番号とともに IGP を用いて広告を行う。その後、同じ Flex Algo に含まれるノードのみが経路計算を行うことで、制約に基づいた最短経路が決定される。Flex Algo はネットワークの柔軟な分割が可能である点から、ネットワークスライシング [16] への利用が期待されている。

図 3.2 に、Filsfils の資料 [17] より Flex Algo の適用例を引用する。この例では、Algo0、Alg128、Alg129 の 3 つの Flex Algo が適用されている。このトポロジにはノード 0 からノード 9 までの 10 個のノードが存在し、全てのノードが Algo0 に参

加している。また、ノード 1・2・3・4 は Alg128 に、ノード 5・6・7・8 は Alg129 に、ノード 0 とノード 9 は Alg128 と Algo129 の双方へ参加している。ここで、各ノードが Algo ごとに Prefix を広告することで、Algo ごとに Node Segment を設定する。例として、ノード 2 は Algo0 には 16002 を、Algo128 には 16802 を広告している。これにより、16002 が指定された場合はトポロジ全体での最短経路が、16802 が指定された場合には、Algo128 の範囲内での最短経路が用いられるようになり、ネットワークスライシングが実現される。

Flex Algo はドメイン分割を目指し考案された技術ではないが、Node SID の広告範囲分割を利用することで、サービスごとのスライシングが可能となる。これを利用することで、複数の AS が存在する環境において、サービスに基づいたネットワークの利用など、柔軟なネットワーク利用を行うことができると考える。ただし、この手法では全てのノードを単一の SR ドメインに含める必要があるため、複数の AS が存在するネットワークへ適用する場合は、管理の分割や設定の分離を行う仕組みが必要となる。

3.2 他の経路制御技術における大規模化手法

本節では他の経路制御技術における大規模化手法について扱い、セグメントルーティングに不足する要件を検討する。

3.2.1 Stateful H-PCE

MPLS におけるドメインを越えた LSP の確立手法として、IETF のワーキンググループである PCE[18] において Stateful Hierarchical Stateful Path Computation Element（H-PCE）[19] が提案されている。H-PCE は MPLS における ACTN を実現する技術として標準化が進められている。

H-PCE の概念を Dhody らの資料 [20] より引用し、図 3.3 に示す。H-PCE では、各ドメインの経路計算を行う子 PCE と、子 PCE を管理する親 PCE が存在する。子 PCE は TED に各ドメイン内の TE 情報を格納し、親 PCE はドメイン単位での隣接関係を管理する。親 PCE は経路計算のコントローラとしての役割を担う。

H-PCE による経路計算手順を以下の（1）から（6）に示す。（1）PCC は経路計算要求を子 PCE へと送信する。（2）子 PCE は TED を参照し、ドメイン内部宛の経路であった場合は子 PCE 自身が経路計算を行う。ドメインを越えた経路であった場合には、親 PCE へ経路計算要求を発行する。（3）親 PCE はドメインの隣接関係情報を元に、ドメイン単位での経路候補を作成する。候補に含まれるドメインの子 PCE へと経路構築を要求する。（4）それぞれの子 PCE は TED とポリシーから経路を構築し、親 PCE へと送信する。（5）親 PCE は送信された経路情報か

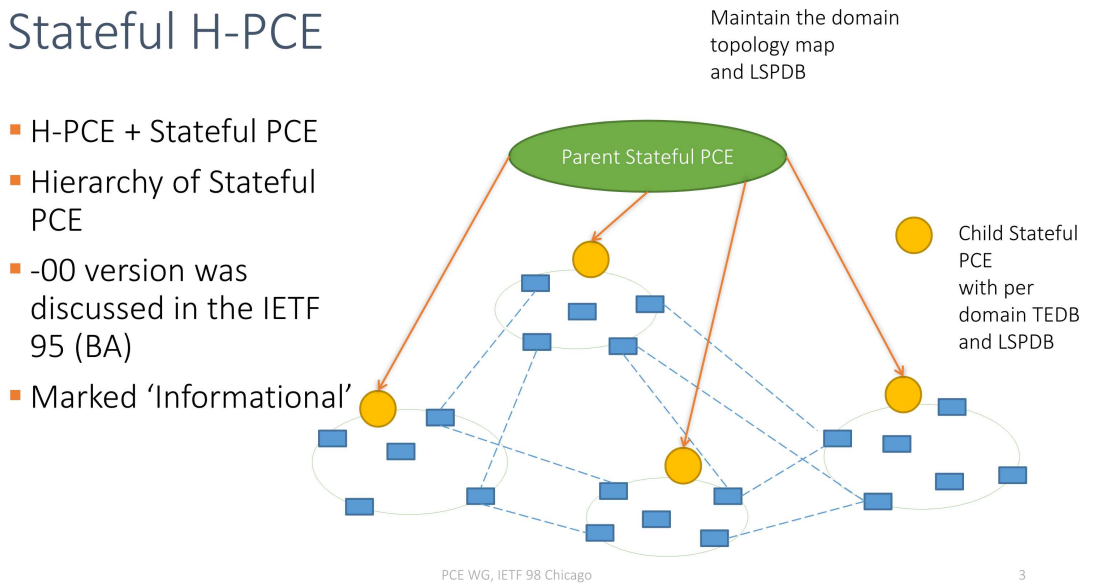

図 3.3: H-PCE（出典：Dhody, et al., Stateful H-PCE & ACTN, 2017.）

ら最適な経路を選択し、一連の経路を生成する。(6) 親 PCE は子 PCE を経由し、PCC へと構築した経路情報を送る。

これらの手順により、ドメインを越えた経路であっても一連の経路計算が可能となる。H-PCE ではドメイン内部の計算は子 PCE が計算すると想定されているが、ISP 内部のホスト同士の通信を外部のパブリッククラウドのネットワーク機能を通過させ行うように、ドメイン内部への経路であっても、ドメイン外を経由した経路の構築を行うことでより柔軟な経路制御が可能となる。本研究ではこの手法についても議論を行う。また、H-PCE では PCC ドメインを越えて宛先ノードの情報を得る手段については考えられていない。この場合、各ノードがドメイン外のノードの情報を得るための仕組みが必要となる。しかしその場合では運用プロトコルの増加による管理の複雑化が生じる上、全てのノードが異なるドメインの情報をノード単位で得ることになる。それに対し、ドメイン全体を管理する主体が通知することにより、AS 単位でのプロトコルの削減や情報の隠蔽が可能となるだろう。本研究ではより柔軟に経路の構築を行うため、経路計算主体の利用についても検討を行う。

第4章　階層型セグメントルーティングの提案

本章では第2章で述べた課題を解決するため、セグメントルーティングに階層構造を導入した階層型セグメントルーティングを提案し、その実現に必要となる機能を議論する。その上で提案手法に対し、従来のセグメントルーティングと等価にパケット転送を行うための実現方式による5つのモデルと、各サブドメインへのポリシー設定方式による2つのモデルを提案し、その特徴と用途について議論を行う。

4.1　セグメントルーティングの階層化

第2章で述べたように既存のセグメントルーティングでは全てのノードが同一のSR domainに参加するため、管理者の異なるネットワークが連携する環境であっても設定やポリシーの共通化が必要となり、管理の分割が困難となる。また全てのノードが情報を共有するため、規模拡大によるネットワークの安定性低下やSIDテーブルのエントリ数が増大する。

これら課題を解決するため、本研究ではセグメントルーティングに階層構造を取り入れた階層型セグメントルーティングを提案する。階層型セグメントルーティングは、複数のノードを束ねたサブドメインを定義することでネットワークをサブドメインごとに分割し、その上でサブドメインを越えた経路制御を実現するためにコントロールプレーンを階層型に構成した上で、経路計算を行う。

既存のセグメントルーティングの概念図を図4.1に、階層型セグメントルーティングの概念図を図4.2に示す。図4.2では、1つのSRドメインの中にA・B・Cの3つのサブドメインを構成した例を示している。階層型セグメントルーティングでは、ポリシー分割・機能分離の明確化のため、SR PCEも同じく階層型に構成する。図4.2ではサブドメインであるA・B・Cに対応したSR PCE a・b・cと、それらを繋いでSRドメイン全体の経路計算を行うSR PCE Σを作成している。ノード23のSIDテーブルには、サブドメインC内部のノード21・ノード22・ノード24のNode SID情報と、ノード23の隣接関係を示す101・102・103のAdjacency SID情報のみが格納されている。それに対して図4.1では既存のセグメントルーティングにより全てのノードが1つのSRドメインの中に属したフラットな構成となって

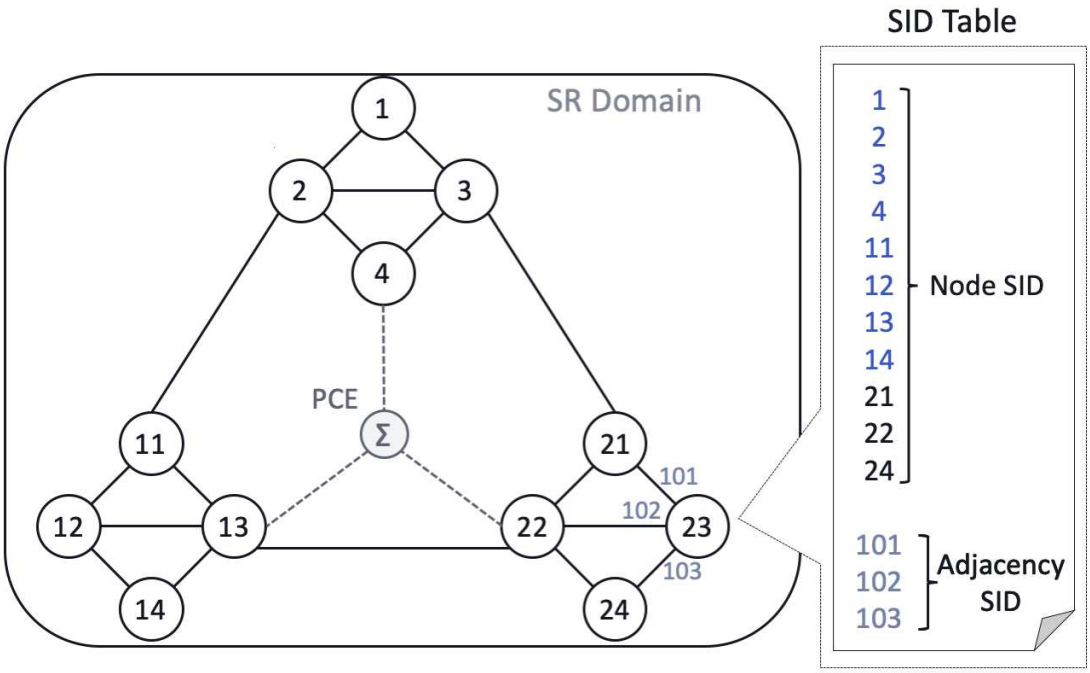

図 4.1: 既存のセグメントルーティングの概念

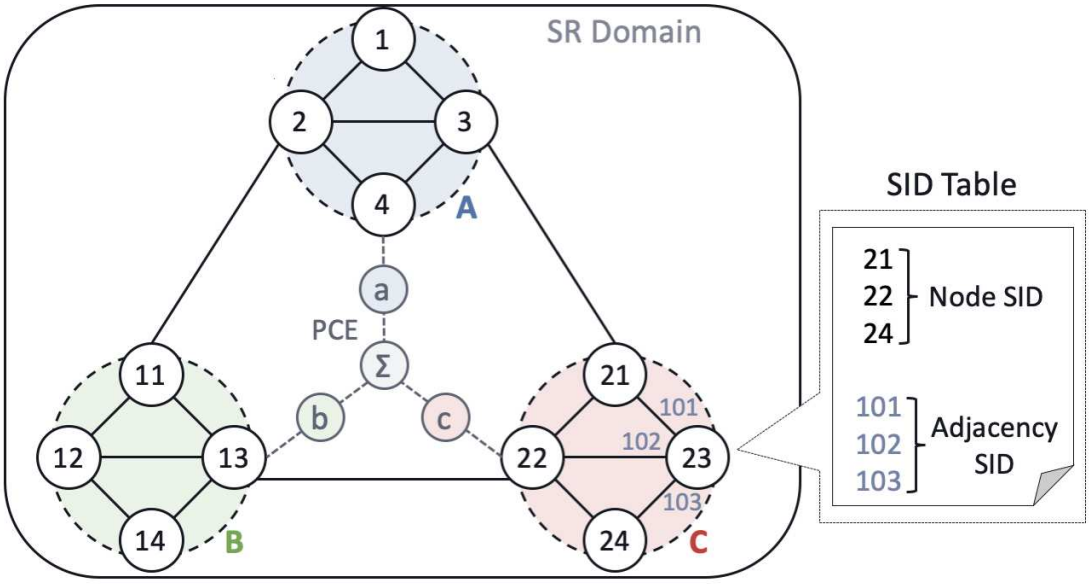

図 4.2: 階層型セグメントルーティングの概念

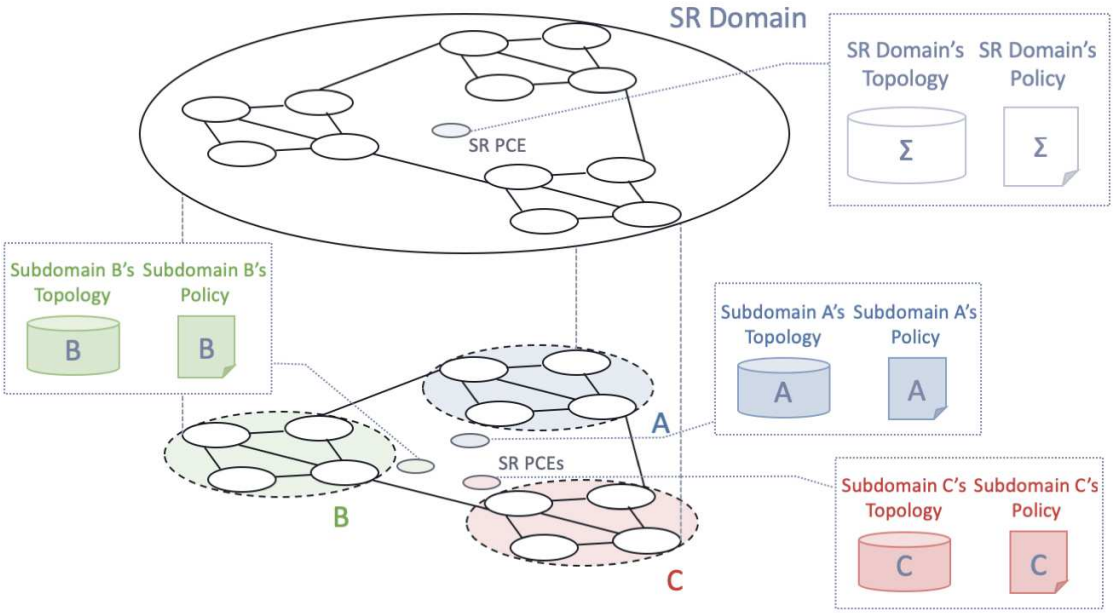

図 4.3: セグメントルーティングの階層化

いる。そのためこの構成ではドメイン内の各ノードが他の全ノードの経路アップデート情報を受け取る必要があり、また SID テーブルエントリも同様に全ノード数を所持する必要が生じる。図 4.1 中のノード 23 は、SID テーブルに他の全てのノードの Node SID 情報と、隣接関係を示す 101・102・103 の Adjacency SID 情報を格納する必要がある。

図 4.2 のように、階層型セグメントルーティングでは AS や拠点等の各ネットワークをサブドメイン単位に分離することで、サブドメインごとに独立した管理構造が実現できる。またそれに伴い、各ノードの保有すべき SID テーブルの対象範囲が縮小し、経路アップデートの影響範囲削減による安定性の向上と各テーブルのエントリ量が削減可能となる。

4.1.1 階層化の定義

本項では本研究における階層化の用語定義を行う。

本研究における階層化とは、セグメントルーティングのコントロールプレーンをサブドメイン分割された下位層とサブドメインを束ねる上位層に分け、管理構造を層状に構成することと定義する。図 4.3 に階層化されたコントロールプレーンの構造を示す。図 4.3 は、第 4.1 節の例と同様に 1 つの SR ドメインを A・B・C の 3 つのサブドメインに分割した例を示している。下位層はサブドメインごとにノードを分割し、計算主体である SR PCE がそれぞれのサブドメインに存在する。

SID テーブルの情報もサブドメインごとに分割するため、各ノードは、それぞれが属するサブドメイン内の SID 情報のみをもち、サブドメイン外の SID 情報は持たない。上位層はドメイン全体を管理する役割をもつ。上位層の SR PCE は SR ドメイン全体を管理する。上位層で全体のトポロジを把握することにより、サブドメインを越えた End-to-End の経路制御を実現できる。H-PCE と異なり、上位層の PCE はドメイン間の繋がりだけでなく、詳細な経路情報を管理する。階層の利用例を以下に述べる。階層を 2 段に構成する場合は、2 つの例が考えられる。1 つ目は AS 間での連携である。下位層を各 AS 等のドメイン単位で構成し、上位層を AS 連携で利用する例である。各 AS が SR PCE を運用し、AS 連携を行う際に契約へ基づいて上位層の SR PCE を設定することで連携を行う。2 つ目は AS 内での管理分割である。コアネットワーク・アクセスネットワーク・データセンタなど、管理分割を行いたい単位でサブドメインを構成し、それらを上位層で連携させる。また、2 つ目の手法で階層化した AS 同士を連携させたり、複数の AS による連携において、さらに特定の AS のみで連携を行う場合などでは、階層を 3 段以上に構成することも可能となる。

このように、サブドメイン分割されたネットワークを階層型に管理することにより、サブドメインごとの情報分割とサブドメインを越えた経路の構築を同時に実現することが可能となる。各層が担う役割の明確化と分割のため、本研究では SR PCE の使用を想定している。階層型セグメントルーティングにおける SR PCE の具体的な構成について、第 4.1.2 項で扱う。

4.1.2 階層型セグメントルーティングにおける SR PCE

図 4.3 では、上位層の SR PCE が全体のトポロジ Σ と、ポリシー Σ を持ち、下位層の SR PCE がそれぞれのサブドメインのトポロジとポリシーを持つ様子を示している。このように、階層型セグメントルーティングでは各下位層の SR PCE がそれぞれのサブドメインのトポロジとポリシーを管理する。上位層の SR PCE は各下位 SR PCE からトポロジを取得することにより、SR ドメイン全体のトポロジを管理する。それに加え SR ドメイン全体へと適用されるポリシーを持つことによりサブドメインを越えた経路計算を実現する。

階層型セグメントルーティングにおける階層型 SR PCE を用いた経路計算の手順を図 4.4 の（1）から（3）に示す。ここで、SR PCE a・b・c はサブドメイン A・B・C を管理する下位 SR PCE であり、SR PCE Σ は SR ドメイン全体を管理する上位 SR PCE である。図 4.4 中の手順（1）では、事前準備として各下位 SR PCE がそれぞれの管理するサブドメイン内部の SID 情報・トポロジ情報を収集する。その後、上位の SR PCE がそれぞれの SR PCE から SR ドメイン全体のセグメント情報・トポロジ情報を収集する。手順（2）では、SR PCE が収集した情報を用い、あらかじめ設定されたポリシーに基づいてセグメントリストを構築する。この例ではノード 23 からノード 1 への経路を示すセグメントリストを構築している。手

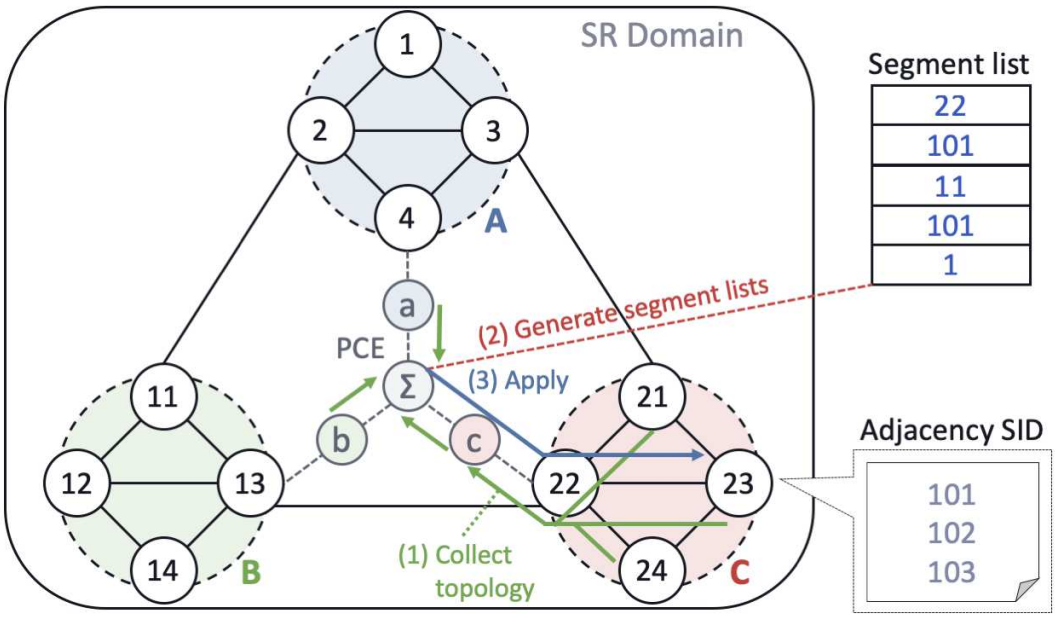

図 4.4: 階層型 SR PCE による経路計算

順（3）では、生成したセグメントリストを各ノードに配布する。この例では手順（2）で生成したセグメントリストをノード 23 へ送信している。このように、階層型セグメントルーティングでは階層化された SR PCE が連携することにより、トポロジの収集と経路計算を実現する。

4.1.3 サブドメインの分割と SID テーブル

階層型セグメントルーティングではサブドメインごとに管理情報の分割を行うため、SID テーブルに格納される情報もサブドメインごとに分割が行われる。そのため、通常のセグメントルーティングと同様の SID 指定ではサブドメインを越えた転送を行うことができない。図 4.5 にサブドメインを越えた指定が行われた様子を示す。

各サブドメインは属するサブドメイン内部の他ノードの Node SID の情報と、各隣接関係を示す Adjacency SID の情報を持つ。図 4.5 の例では、ノード 12 はサブドメイン B に属するノード 11・ノード 13・ノード 14 の Node SID と隣接関係を示す 101・102・103 の Adjacency SID を、ノード 23 はサブドメイン C に属するノード 21・ノード 22・ノード 24 の Node SID と隣接関係を示す 101・102・103 の Adjacency SID を SID テーブルに格納している。ここで、ノード 23 からノード 1 へ、ノード 12 とノード 3 を経由するようにセグメントリストを構築することを考える。この場合、通常のセグメントルーティングでは<12, 3, 1>というセグメントリストが構築される。しかし、階層型セグメントルーティングでは SID テーブル

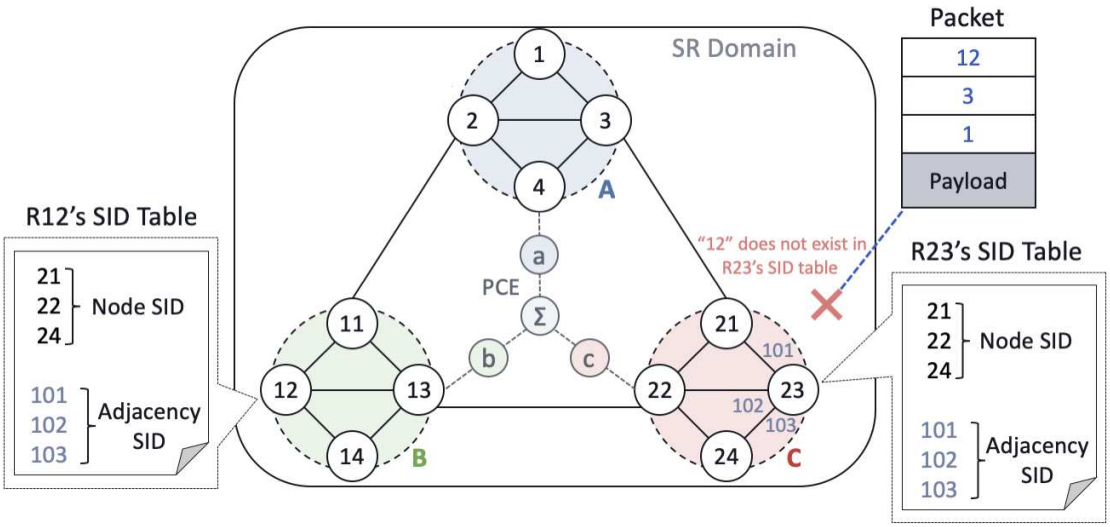

図 4.5: サブドメインを越えた指定の問題

の情報はサブドメイン内部に限定されているため、ノード 23 の SID テーブルにはノード 12 の Node SID は含まれていない。そのため作成したセグメントリストによる転送を行うことが不可能である。この課題を解決するため、第 4.2 節でサブドメイン越えの手法を提案しその特徴によりモデル化を行う。

4.2 サブドメインを越えた経路制御の手法によるモデル化

第 4.1.3 項で、階層型セグメントルーティングでのサブドメイン越えにおける課題について述べた。本研究の目的である複数の AS が存在する環境におけるサービスチェイニングを実現するためには End-to-End での経路構築・転送が不可欠であり、サブドメインを越えるための手法が必要となる。そこで本研究では、サブドメインを越えた転送を行うための手法によりモデル化を行い、それぞれの特徴の整理と用途ごとの有効な手法を検討する。

4.2.1 展開モデル

展開モデルでは、サブドメインを越える経路においてサブドメインのエッジノードの Adjacency Segment を利用することにより転送を実現する。展開モデルの概念図を図 4.6 に示す。

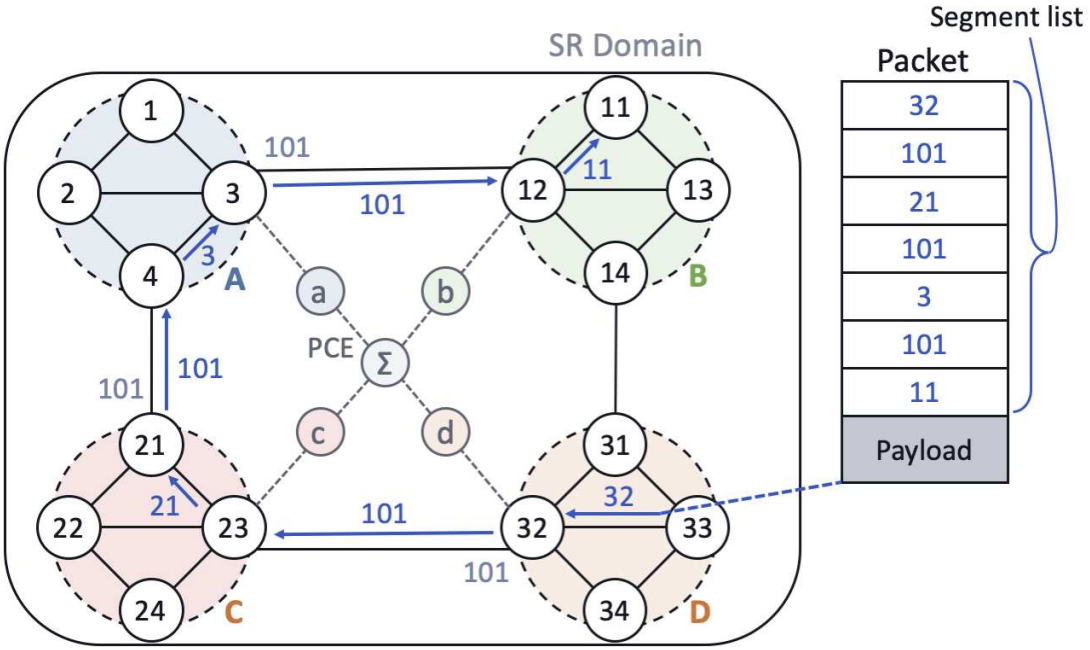

図 4.6: 展開モデル

展開モデルでは、各ノードはSIDテーブルに自らの属するサブドメイン内のNode SIDと隣接関係のAdjacency SIDを持つ。セグメントリスト構築の際にサブドメインを越えるようなNode Segmentを指定した場合、サブドメイン越え経路をサブドメインのエッジノードのNode Segmentとサブドメインを越えるためのAdjacency SIDの組み合わせに展開することで、転送可能なセグメントリストへと変換を行う。図4.6にサブドメインC内のノード21を通過するようにノード33からノード11への経路を構成する例を示した。また比較のため、図4.7に既存のセグメントルーティングにより転送を行なった様子を示す。

図4.7のように、既存のセグメントルーティングでは<21, 11>というセグメントリストを指定することにより、目的の経路を満たす転送を実現できる。一方、展開モデルを利用した例では、サブドメインDからCへ、CからAへとサブドメインを2度越えるため、2度の展開を行う。ノード33からNode SID 21への指定は、エッジノードであるノード32のNode SID 32と、ノード32からノード23へのAdjacency SID 101の2つのSIDへと展開される。同様にノード21からNode SID 11への指定はノード21からノード4へのAdjacency SID 101とNode SID3、ノード3からノード12へのAdjacency SID 101へと展開される。このため展開モデルでは<32, 101, 21, 101, 3, 101, 11>というセグメントリストが構築され、転送が行われる。

このように、展開モデルに従いセグメントリスト内のサブドメインを越えた指定を展開することで、階層型セグメントルーティングにおいて各ノードが持つ情

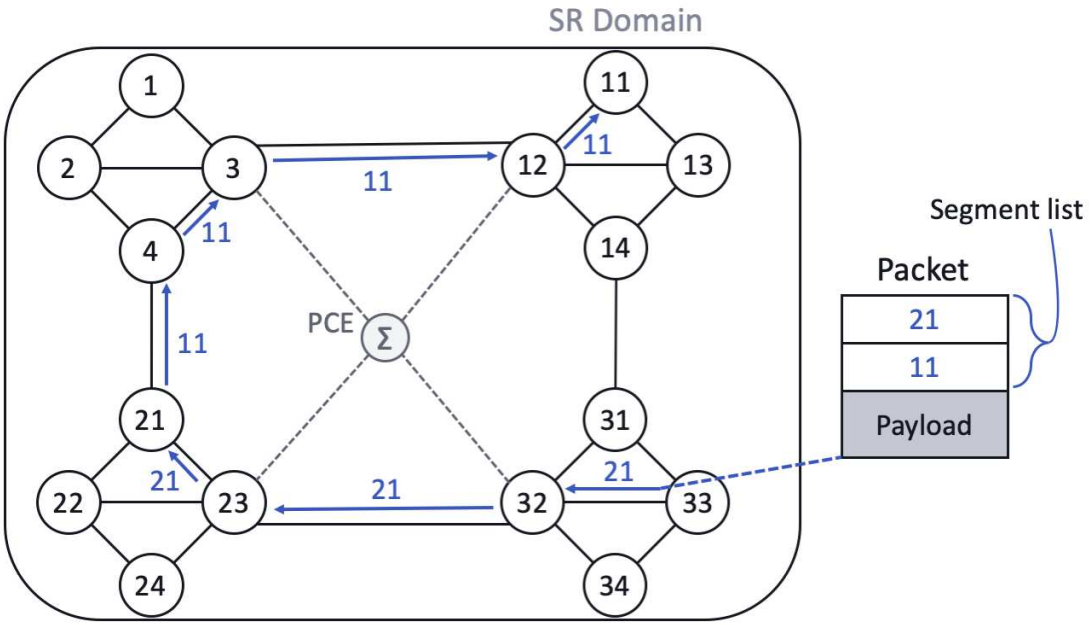

図 4.7: 既存のセグメントルーティングによる転送

報のみで転送可能なようにセグメントリストを構築できる。

展開モデルの利点として、以降で扱う他モデルと比べ、Node SID と Adjacency SID のみを利用するためシンプルな実装が可能となる。その一方で、サブドメインを越える SID を展開するため、上位 Node SID モデルや上位 Adjacency SID モデル等の他モデルと比べセグメントリストの増大が生じる。

4.2.2 上位 Node SID モデル

上位 Node SID モデルでは、各サブドメインに SR ドメイン全体で一意な上位 Segment を作成し、Node SID と同様に上位 Node SID を転送対象として扱う。上位 Node SID モデルの概念図を図 4.8 に示す。

上位 Node SID モデルでは SR ドメイン内の各サブドメインを示す上位 Segment を構成する。各ノードは自らの属するサブドメイン内の Node SID と、Adjacency SID の他に、他のサブドメインを示す上位 Node SID の情報を持つ。各上位 Segment は Node Segment と同様にサブドメインへの最短経路を示す。図 4.8 の例では、サブドメイン D に属するノード 33 は、サブドメイン D 内のノード 31・ノード 32・ノード 34 の Node SID と隣接関係を示す 101・102・103 の Adjacency SID の他に、サブドメイン A・B・C を示す上位 Node SID を SID テーブルに登録する。上位 Node SID の実現手法としては Binding Segment の利用を想定する。

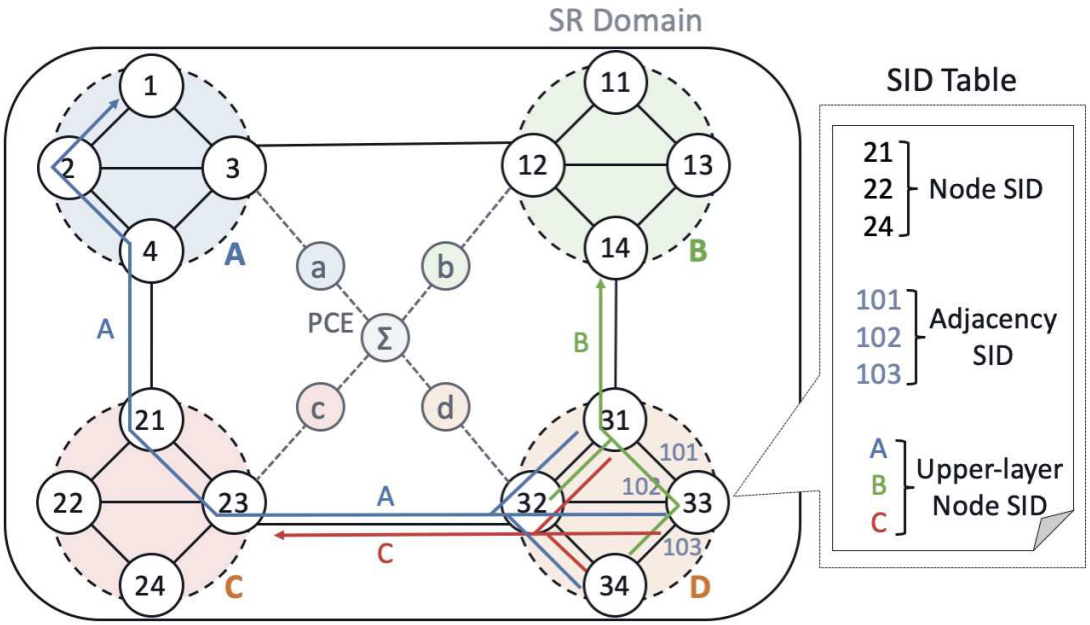

図 4.8: 上位 Node SID モデル

上位 Node SID モデルでは各ノードが他サブドメインを示す上位 Node SID を持つことにより、サブドメインへの転送が可能となる。図 4.9 に、上位 Node SID モデルを用いノード 33 からノード 11 へノード 21 を経由するように転送を行う例を示す。

この例でのセグメントリストは<C, 21, B, 11>となる。ノード 33 は SID テーブルを参照し、上位 Node SID C が示すインターフェースから送出を行う。ノード 32 は SID テーブルを参照し、セグメントリストの先頭要素を次に進め、上位 Node SID C が示すインターフェースから送出を行う。ノード 23 は SID テーブルを参照し、セグメントリストの先頭要素を次へ進めノード 21 へと転送を行う。ノード 21 は SID テーブルを参照し、上位 Node SID B が示すインターフェースから送出を行う。ノード 4 も同様に上位 Node SID B へ送出を行う。ノード 3 はセグメントリストの先頭要素を次へ進め、上位 Node SID B が示すインターフェースから送出を行う。ノード 12 が SID テーブルを参照し、セグメントリストを除去しノード 11 へと転送を行うことでセグメントルーティングによる転送が終了する。

上位 Node SID モデルでは各サブドメインを直接転送先に指定可能であるため、展開モデルに比べてセグメントリストの長さを削減できる。しかし上位 Node SID モデルの課題として、各 BSID を全ノードで生成・広告する仕組みが必要となる。上位 Node SID モデルでは直接接続していないサブドメインの上位 Node SID も登録する必要があるため、ルータのみを用いた SID の共有では、EGP と IGP の双方を利用する必要がある。その対策として、全体のトポロジを管理する上位 PCE

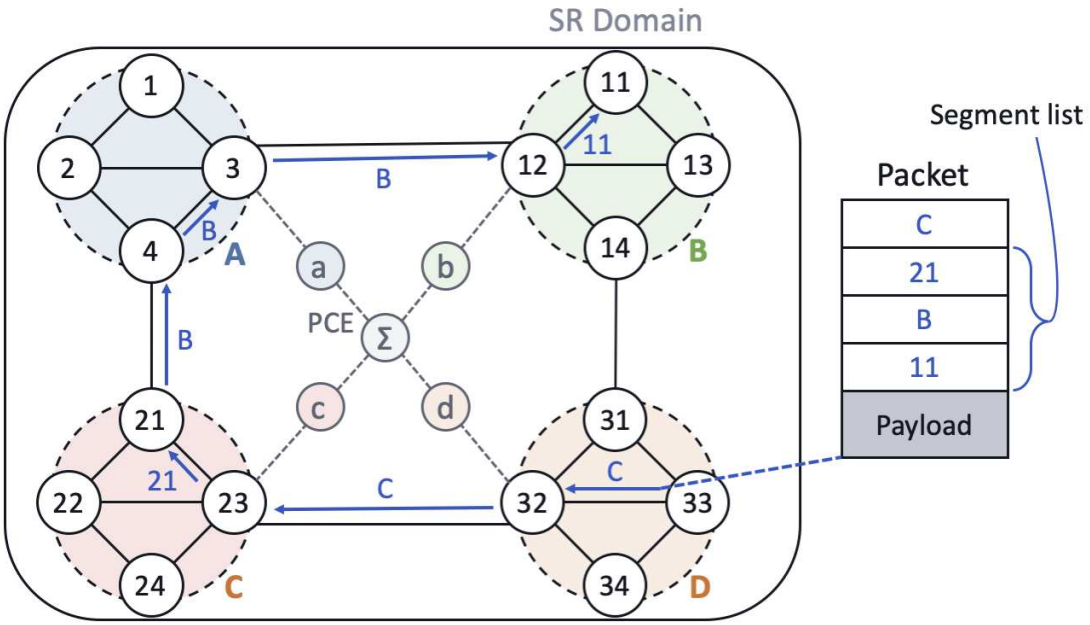

図 4.9: 上位 Node SID モデルによる転送

から各ノードへと設定を行うことが考えられる。

4.2.3 上位 Adjacency SID モデル

上位 Node SID モデルでは、直接接続していない上位 Node SID を把握する必要がある。そのため SID テーブルに全サブドメインの上位 Node SID 情報を登録する必要が生じ、経路更新の影響範囲が広がることでネットワークの安定性が低下する。これに対し、他のサブドメインに SID を付与せず、サブドメインの各隣接関係に SID を付与するモデルを上位 Adjacency SID モデルと呼ぶ。上位 Adjacency SID モデルの概念図を図 4.10 に示す。

上位 Adjacency SID モデルでは、各ノードは自らの属するサブドメイン内の Node SID と隣接を示す Adjacency SID の他に、隣接するサブドメインへの経路を示す上位 Adjacency SID を持つ。図 4.10 の例では、ノード 33 は隣接するサブドメインであるサブドメイン B・C の上位 Node SID を保有しており、隣接していないサブドメイン A の上位 Node SID は保有していない。

図 4.11 に、上位 Adjacency SID モデルを利用しノード 33 からノード 11 へ、ノード 21 を経由するように転送を行う例を示す。この例でのセグメントリストは<C, 21, A, B, 11>となる。ノード 33 は SID テーブルを参照し、上位 Adjacency SID C が示すインターフェースから送出を行う。ノード 32 は SID テーブルを参照し、セグメントリストの先頭要素を次へ進め、上位 Adjacency SID C が示すインター

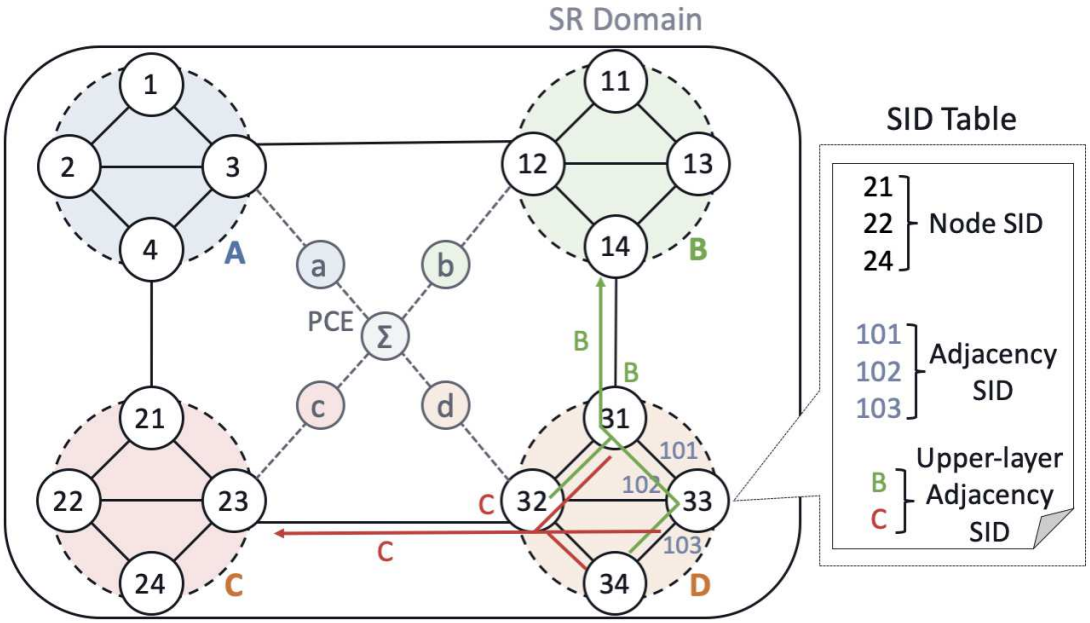

図 4.10: 上位 Adjacency SID モデル

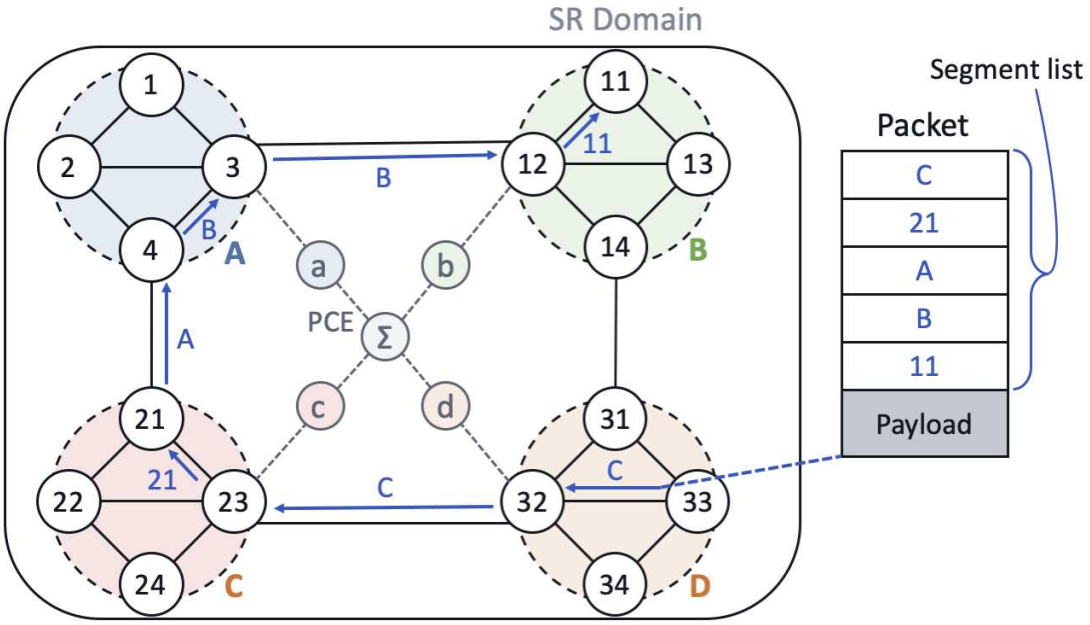

図 4.11: 上位 Adjacency SID モデルによる転送

フェースから送出を行う。ノード 23 は SID テーブルを参照し、セグメントリストの先頭を次へ進め、先頭要素である 21 が示すインターフェースから送出を行う。ノード 21 は SID テーブルを参照し、セグメントリストの先頭要素を次へ進め、上位 Adjacency SID A が示すインターフェースから送出を行う。ノード 4 は先頭要素である B が示すインターフェースから送出を行う。ノード 3 が SID テーブルを参照し、セグメントリストの先頭要素を進め、上位 Adjacency SID B が示すインターフェースへと送出する。ノード 12 がセグメントリストを除去しノード 11 へと転送を行うことでセグメントルーティングによる転送が終了する。

上位 Adjacency SID モデルでは直接隣接していないサブドメインを示す上位 SID を持たないため、経路更新の影響を隣接サブドメインまでに留めることができ、上位 Node SID モデルに比べネットワークの安定性が向上する。しかし欠点として、直接隣接していないサブドメインに対しての転送は隣接するサブドメインへの上位 Adjacency SID を用い中継することによって転送を行うため、セグメントリストの長さが上位 Node SID モデルに比べて増大する。

4.2.4 経路 SID モデル

展開モデルや上位 Node SID モデル、上位 Adjacency SID モデルではサブドメインを越える経路が生じた場合、既存のセグメントルーティングに比べセグメントリストの肥大が生じる。その場合、Maximum Transmission Unit（MTU）の制限により運送可能となるペイロードの容量が制限されるため、セグメントリスト長は縮小することが望ましい。

これを実現するための手法として、経路 SID モデルと再発行モデルの 2 種が考えられる。経路 SID モデルは、セグメントルーティングにより制御を行う経路をセグメントとし、SID を発行する手法である。この手法では、サブドメインを越えた経路が発行された際、その経路を示す経路 SID を発行し、経由地全てに適用する。

図 4.12 に、経路 SID モデルの概要を示す。以降の例では、経路 SID は 33-1 のように送信元と送信先を並べて示す。ここではノード 33 からノード 1 とノード 33 からノード 11 へ、ノード 21 を経由するように経路 SID を適用した例を示した。経路の中継点となるノードには各経路 SID が SID テーブルに含まれている。

図 4.13 に、経路 SID モデルを利用しノード 33 からノード 11 へ、ノード 21 を経由するように転送を行う例を示す。経路 SID モデルでは経路に SID が発行されているため、セグメントリストにはその SID のみが指定される。中継する全てのノードにおいて、先頭要素である 33-11 を参照することで転送が行われる。

経路 SID モデルの利点として、経路を SID で表現するためセグメントリストを 1 つのセグメントで構成できる。そのため他の階層型セグメントルーティングのモデルや一般的な SR と比較し、セグメントリストの削減が可能となる。ただし経路 SID モデルの欠点として、同じノードを 2 度経由する経路を表現できないため、複

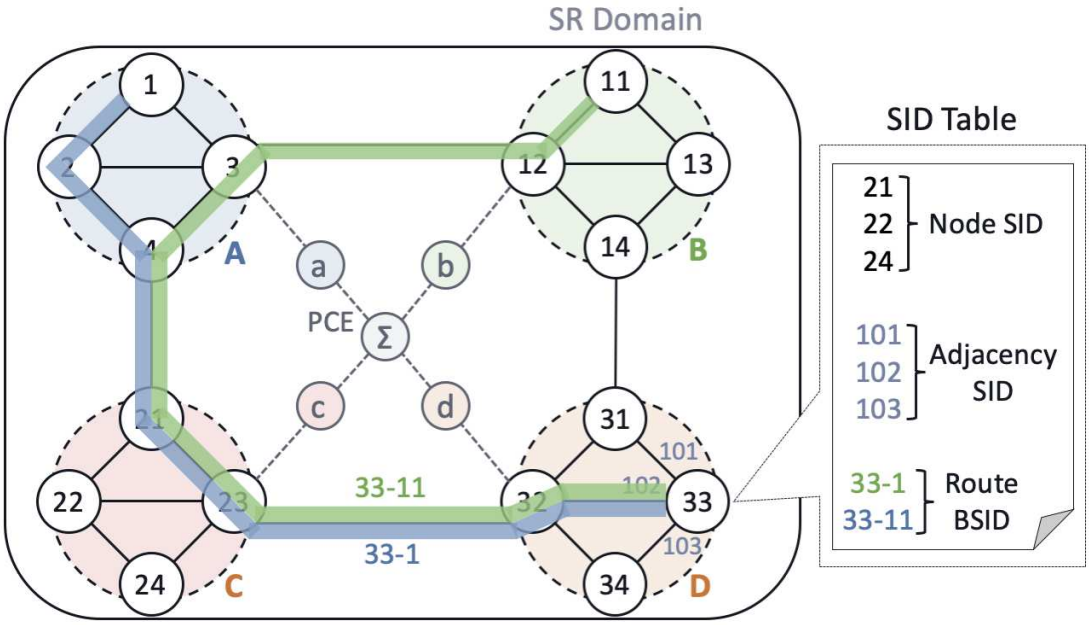

図 4.12: 経路 SID モデル

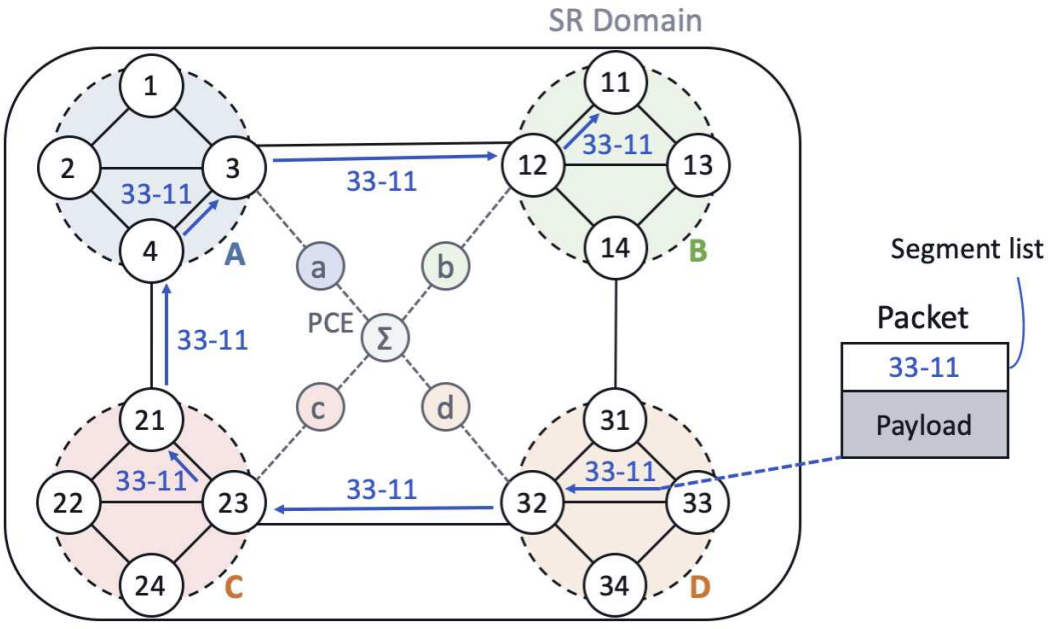

図 4.13: 経路 SID モデルによる転送

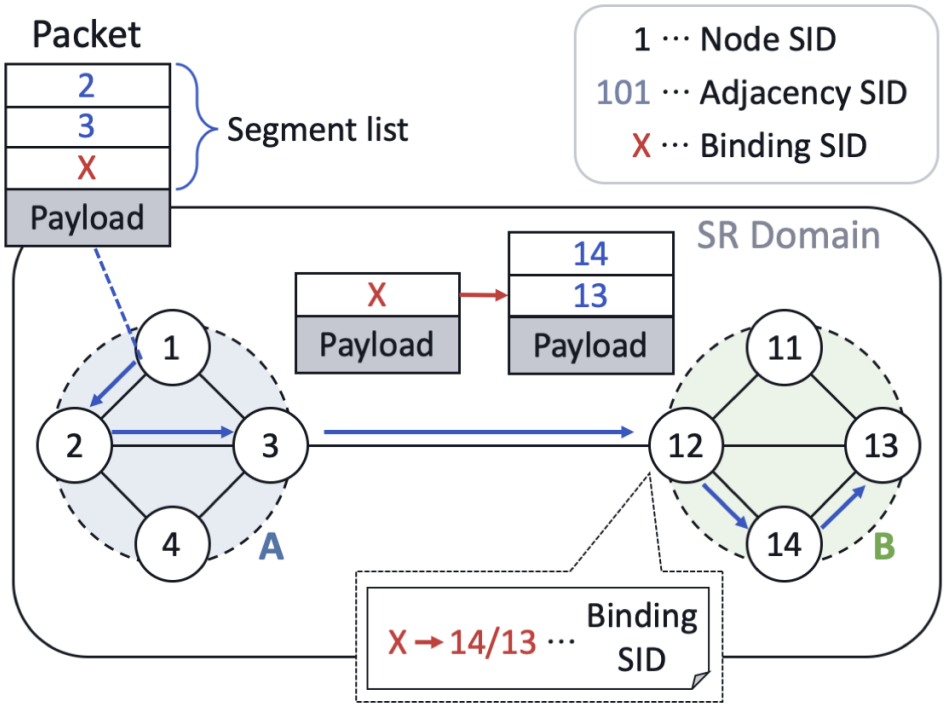

図 4.14: 再発行モデル

数の経路 SID に分割を行うか、Node Segment や Adjacency Segment など他の情報と組み合わせる必要がある。また、経路が新たに発行されるごとにドメインを越えて SID テーブルが更新されることにより、ネットワークの安定性低下と SID テーブルの増大が生じる。それに加え、管理者の異なるネットワークにおいて、サブドメインのエッジルータへ発行された Binding SID を反映させる仕組みが必要となり、RSVP-TE におけるシグナリングと同様に、全ての経由ノードに経路 SID の設定処理が必要となる。BSID の設定手法としては、階層型 SR PCE からの指示や IGP、EGP による配布が考えられる。

4.2.5 再発行モデル

再発行モデルはセグメントリストをサブドメインで完結し、BSID を利用してサブドメインごとに付け直す手法である。

図 4.14 のように、サブドメインを越えた経路を示すセグメントリストには、送信元ノードと同じサブドメインの情報で構成されたセグメントリストと、他サブドメインの経路を示す BSID が含まれる。サブドメイン間の通信ではセグメントリストに識別子となる BSID が先頭要素となる。例では BSID である X を受信したノード 12 は、X に対応するセグメントリスト<14, 13>へと変換を行い、その先の転送を実現する。再発行モデルでは、サブドメインによる分割とセグメントリ

表 4.1: サブドメイン越えモデルの特徴

	既存手法	展開モデル	上位 Node SID モデル	上位 Adj. SID モデル	経路 SID モデル	再発行モデル
セグメントリスト	短	長	短	中	最短	短
AS 連携時の管理コスト	不可	小	小	小	大	中
実装負荷	無	小	大	大	大	中

ストの削減を同時に実現できる。第 3.1.1 項で扱った SPRING による拠点間通信の転送手法はこのモデルに分類される。再発行モデルの課題として、サブドメインのエッジとなるノードへ BSID の適用が必要となる。BSID 生成・広告に伴う経路表の更新により展開モデルや上位 SID を利用するモデルに比べネットワークの安定性が低下する。BSID 設定の手法としては、経路 SID と同じく階層型 SR PCE からの配布や IGP、EGP による配布が考えられる。

4.2.6 サブドメイン越えの手法による各モデルの比較

各モデルの特徴を表 4.1 に比較する。モデルごとに利点と欠点が存在し特徴も異なるため、階層型セグメントルーティングの運用では、ネットワークの構成と特性に応じて各モデルの使い分けが必要である。

具体的な使い分けとして、運用者の管理コストと PCE の実装負荷を考慮した際には Node Segment と Adjacency Segment のみで実現できる展開モデルが適している。しかし展開モデルにはセグメントリスト肥大化の欠点が存在するため、パケットのペイロード長と管理コスト・PCE の実装負荷を両立する場合は上位 Node SID モデル、あるいは上位 Adjacency SID モデルが適している。これら 2 つの違いとしては、上位 Node SID モデルはセグメントリストを圧縮できるがネットワークの安定性が下がり、上位 Adjacency SID モデルはセグメントリスト長が増える代わりにネットワークが安定する。

ペイロード長が重要になる環境においては、再発行モデルや経路 SID モデルが適している。ただしどちらのモデルも新たな経路の設定時に BSID の配布処理が生じる可能性がある。

4.3 ポリシー設定手法によるモデル化

階層型セグメントルーティングでは各サブドメインごとに管理者が異なるネットワークに対する利用を想定する。管理者の異なるネットワークではポリシーも

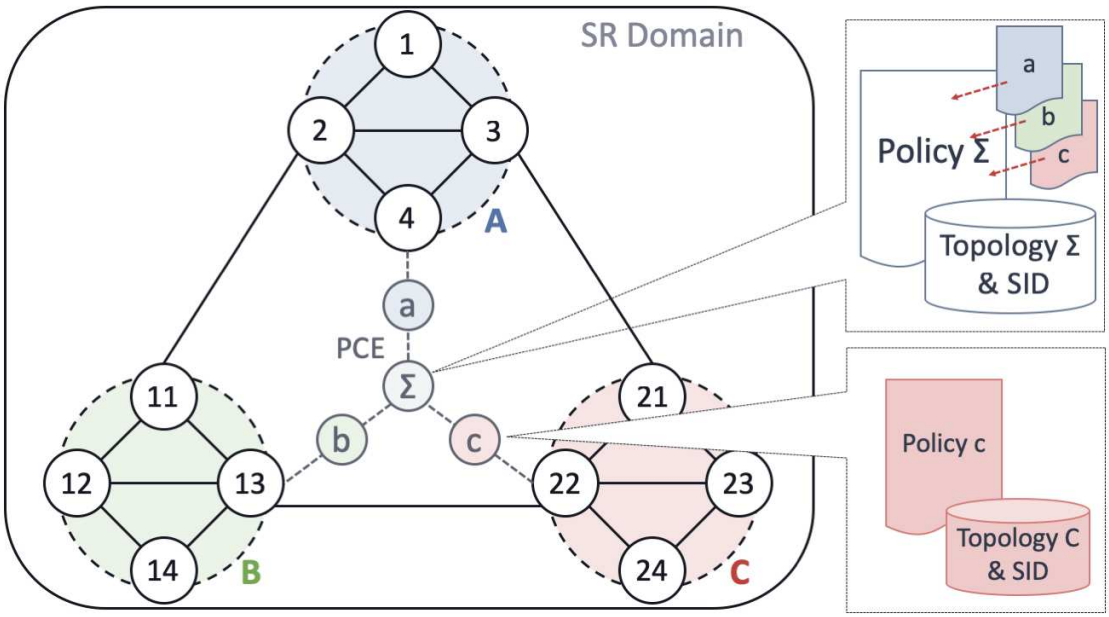

図 4.15: 収集モデル

ネットワークごとに異なるため、各サブドメインや SR ドメイン全体に対するポリシーの適用範囲と適用手法が重要となる。本研究では、階層型セグメントルーティングにおけるポリシー設定手法について、収集モデルと配布モデルの 2 種類のモデルを提案する。

4.3.1 収集モデル

収集モデルは帯域制限等の QoS やスライス分割等のポリシーを各サブドメインごとに設定し、それらを合わせることで上位 SR PCE が用いる SR ドメイン全体のポリシーを作成する方式である。収集モデルでは管理者が異なるネットワークをサブドメインとし、階層型セグメントルーティングを使用することを想定する。そのような適用例では各管理者が他の管理者とは独立にネットワークのポリシーを作成する。各管理者が作成したポリシーを下位 SR PCE に設定し、上位 SR PCE が下位 SR PCE からポリシーを収集することで SR ドメイン全体のポリシーを作成する。

図 4.15 に、下位 SR PCE である a・b・c のポリシーを収集し、SR ドメイン全体のポリシーを Σ に構築した様子を示した。この例では、サブドメイン C の下位 SR PCE にポリシーが設定されている。SR ドメイン全体を管理する上位 SR PCE には、a・b・c それぞれの下位 PCE に適用されたポリシーが集められ、全体のポリシーが作成される。

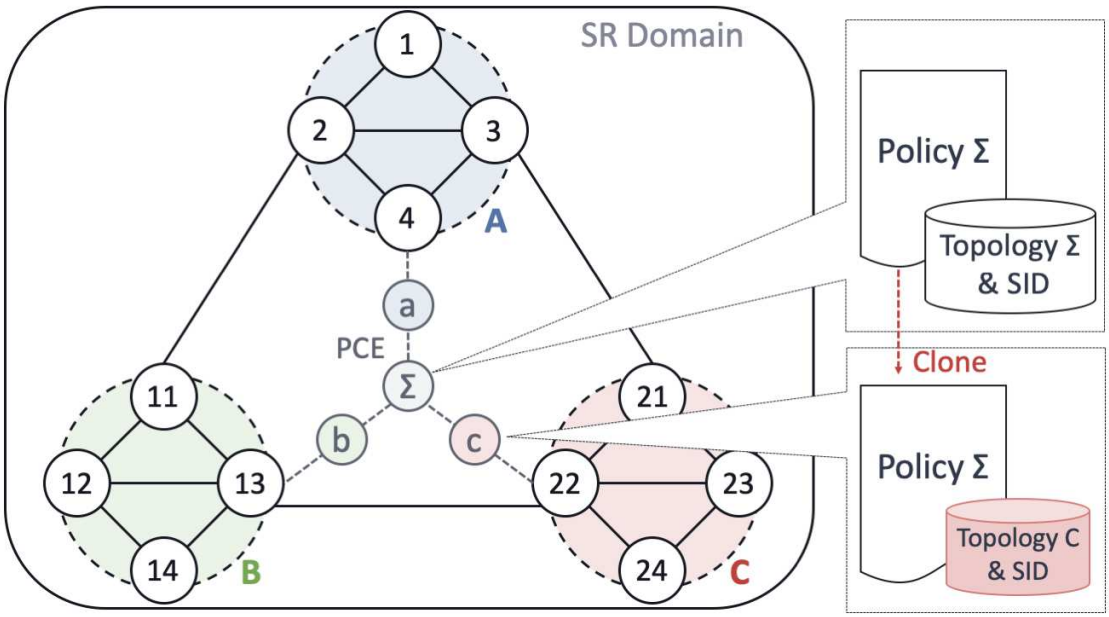

図 4.16: 配布モデル

収集モデルの特徴として、サブドメインごとのポリシーの分離が容易であり、異なるポリシーで運用されるネットワーク間の連携に適する点が挙げられる。

4.3.2 配布モデル

配布モデルは、SR ドメイン全体で統一されたポリシーを作成し、上位 SR PCE で全サブドメインのポリシーを一元管理する手法である。配布モデルの概念図を図 4.16 に示す。

配布モデルでは、管理者が同一のネットワークが地理的・管理的要因で運用の切り離されている環境で、階層型セグメントルーティングにより一連の経路制御を行うような利用形態を想定する。例として、単一の管理者の所有するデータセンタが複数の国に分散し配置された環境が想定される。このような場合では、各データセンタネットワークは地理的に離れており、それぞれが独立し運用される。しかし管理者が同一であるため、ポリシー決定・管理を一元的に行うことが可能である。配布モデルでは、ネットワークの管理者が上位の SR PCE にのみポリシーを設定し、上位 SR PCE が必要な設定のみを下位の各 SR PCE へと配布することで、SR ドメイン全体にポリシーを設定する。図 4.16 では、中央の上位 SR PCE Σ に設定されたポリシーが各下位 SR PCE に配布されることにより、SR ドメイン全体にポリシーが作成されている。

配布モデルの特徴として、ポリシーの一元管理が可能な他、上位・下位の SR PCE を一体として扱いやすい点が挙げられる。

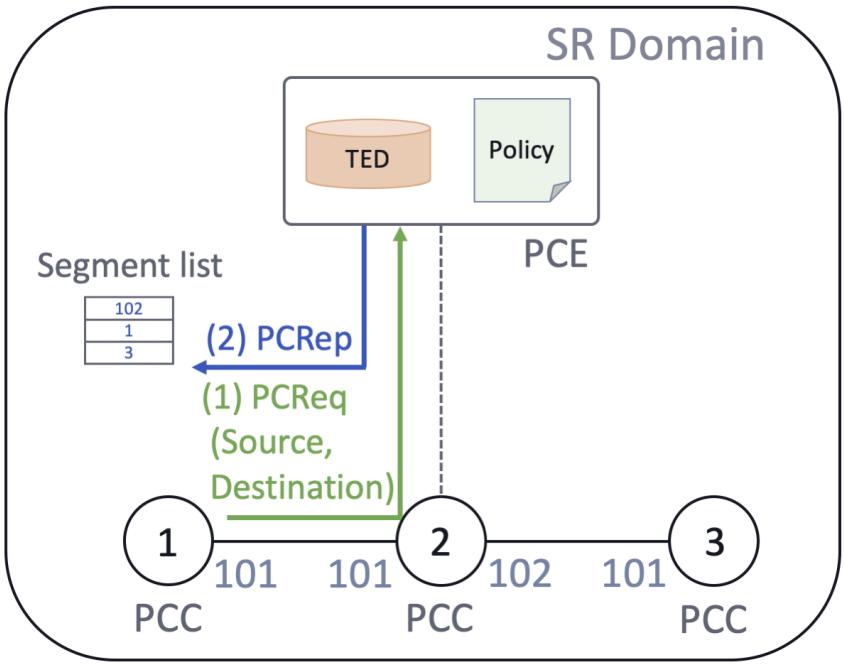

図 4.17: pull 型の経路計算要求

4.3.3 ポリシー設定手法による各モデルの比較

ポリシー設定手法による 2 つのモデルは、それぞれ利用すべきシナリオが異なる。収集モデルはサブドメインごとにポリシー設定ができる性質上、AS 連携によるサービスチェイニング提供など、運営者の異なるネットワーク同士のポリシーを分離した連携に適する。それに対し配布モデルは統一されたポリシーの配布を行う性質や階層化された SR PCE を一元管理可能であり、一体として構築できるという性質上、複数拠点のデータセンター等の、運営者が同一であるが地理的に離れた拠点の連携に適している。

4.4 経路計算要求の手法による SR PCE のモデル化

セグメントリスト構築処理を設計する際、経路計算要求の方式により複数の手法が考えられる。本研究では経路計算要求の方式を push 型と pull 型の 2 種に分類する。

図 4.17・図 4.18 にリクエスト方式の 2 種類のモデルを示す。

1 つ目の設計手法は pull 型である。このモデルでは、PCC は宛先と送信元の組み合わせを PCReq に含め送信する。PCE は PCReq で指定された宛先と送信元に対応したセグメントリストを構築し、PCRep として配付を行う。図 4.17 では、ノード 1 が PCE へ送信元と宛先の組み合わせを送信している。PCE はポリシーの中か

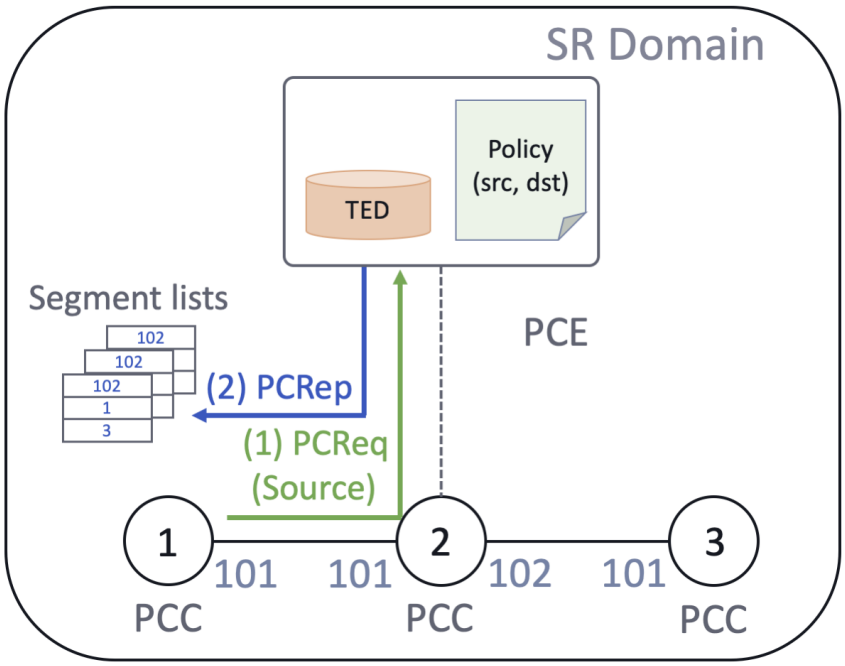

図 4.18: push 型の経路計算要求

ら送信元と宛先の組み合わせに一致するものを選択し、セグメントリストを構築、PCRep としてノード 1 へと送信している。2 つ目の設計手法は push 型である。このモデルでは、PCC は送信元を指定した PCReq を送信する。PCE は送信元と宛先のリストを保有しており、送信元に対応する宛先へのセグメントリストを構築し、PCRep として送信する。図 4.18 では、ノード 1 が PCE へ送信元を指定した PCReq を送信し、PCE はポリシーの中から送信元と一致する組み合わせを検索し全てのセグメントリストを構築、PCRep としてノード 1 へと送信している。

第 3.2.1 項で述べた H-PCE は、pull 型に分類できる。pull 型の利点として、1 つのリクエストが 1 つの経路計算要求に対応しているため、PCC が必要な際に経路を要求できる点が挙げられる。ただし、pull 型では各 PCC が他のドメインに存在するノードを知るための仕組みが必要となる。それに対し push 型では、発行する全経路を PCE で一元管理可能である点が利点として挙げられる。これにより、各 PCC が宛先となるノードの情報を保持していなくてもセグメントリストを発行でき、また各ノードが他ドメインに存在するノードの情報を得るための手法が不要となる。そのため、本研究の提案手法である階層型セグメントルーティングには push 型が適している。

第5章　階層型セグメントルーティングの設計と実装

本章では、提案手法である階層型セグメントルーティングの設計と実装について述べる。

5.1　階層型セグメントルーティングの設計

本節では階層型セグメントルーティングに必要な機能を整理し設計を行う。設計は経路計算主体である階層型 SR PCE の設計と他の要素を含めた階層型セグメントルーティング全体の構成に分けて行う。

5.1.1　階層型 SR PCE の設計

階層型セグメントルーティングを実現するため、経路計算主体となる SR PCE の設計を行う。

既存の SR PCE の設計を図 5.1 に示す。図の下部が経路計算を依頼する SR PCC の処理であり、上部が経路計算を実行する SR PCE の処理である。また、図の左側に TED の更新処理を、右側にセグメントリスト構築処理を示した。

TED の更新処理では、SR PCC がトポロジ情報・トラフィックエンジニアリング情報を含むリンクステート情報を送出する。SR PCE 側ではリンクステートを受信し TED へ格納する。

セグメントリスト構築処理では、まず SR PCC が経路計算要求である PCReq を送信する。階層型セグメントルーティングでは各ノードがサブドメイン外のノードの情報を持たないことから、SR PCC が宛先の情報を持つ pull 型ではなく、SR PCE が送信元と宛先の情報を管理し、SR PCC は送信元のみを含んだリクエストを発行する push 型のモデルを採用する。送信元が指定された PCReq を受信した SR PCE は TED とポリシーを参照し、経路制御に必要なリンクステートとポリシー、宛先の情報を取得する。取得した情報を元に CSPF 計算を行うことで制約付き経路が構成され、経路に合わせセグメントリストを構築し SR PCC へと送信を行う。SR PCC ではセグメントリストを受信し、スタティックルートとして経路表に登録することで経路構築が完了する。

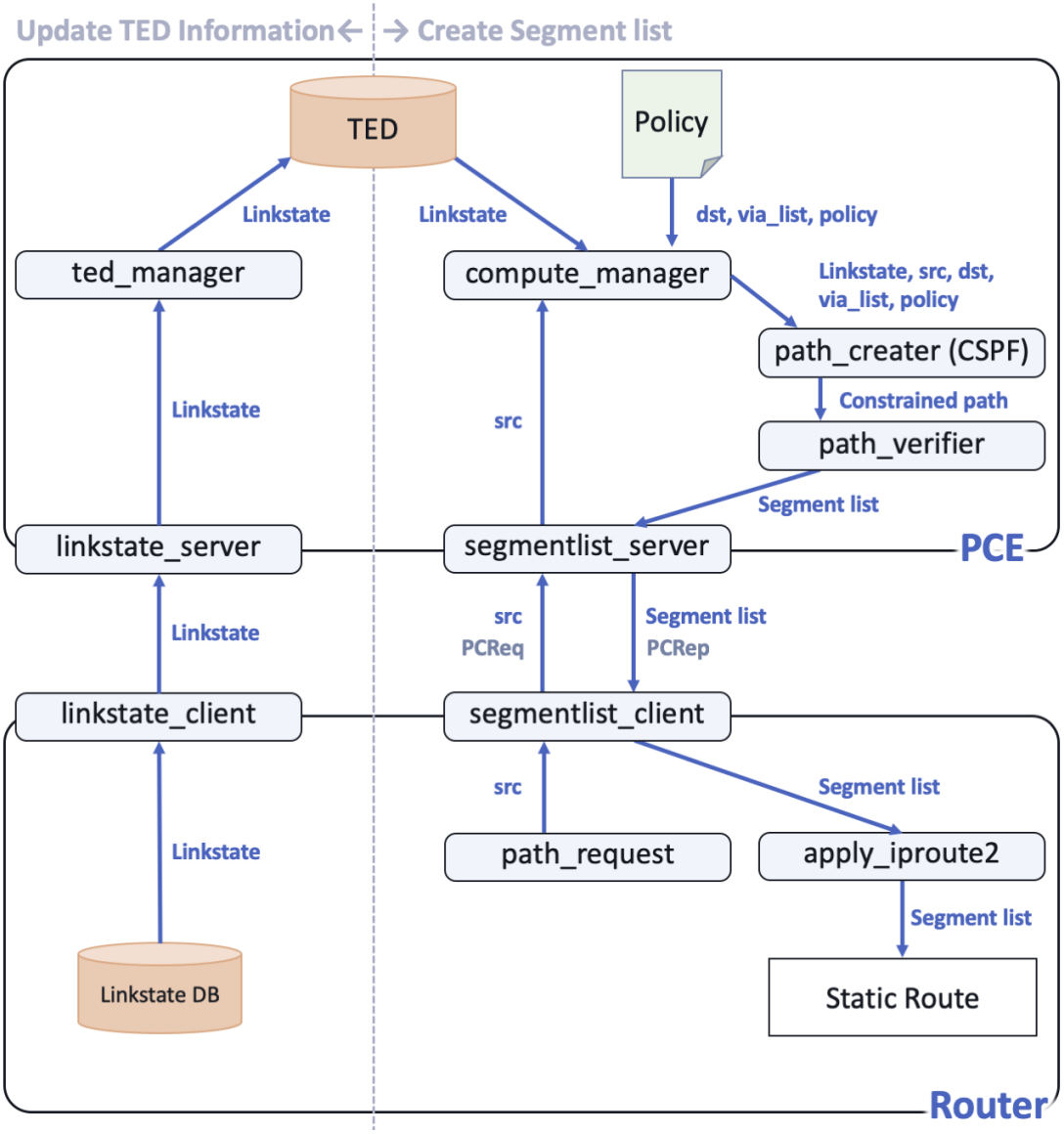

図 5.1: 既存 SR PCE の設計

既存の SR PCE を元に階層型 SR PCE の設計を行う。図 5.2 に階層型 SR PCE の設計を示す。

階層型 SR PCE では、既存の SR PCE と同様に構成する上位 SR PCE と既存の SR PCC の中間に、新たな処理を追加した下位 SR PCE を構成する。図の左側に示した TED の更新処理では、SR PCC は既存の SR PCE と同様にリンクステートを送信する。中間層である下位 SR PCE では受信したリンクステートを TED に格納し、更に上位 SR PCE へと送信を行う。上位 SR PCE がリンクステートを TED へ格納することで、全ての TED 更新処理が完了する。

セグメントリスト構築処理では、まず SR PCC が既存の SR PCC と同様に PCReq を送信する。送信元が指定された PCReq を受信した下位 SR PCE は上位 SR PCE への PCReq の転送とセグメントリスト構築の両方を行う。下位 SR PCE から PCReq を受信した上位 SR PCE は、既存の SR PCE と同様にセグメントリスト構築を行い、下位 SR PCE へ送信する。下位 SR PCE は自らの計算したセグメントリストと上位 SR PCE から受け取ったセグメントリストをそれぞれ SR PCC に送信する。SR PCC が受信したセグメントリストをスタティックルートとして経路表へ登録することで経路構築が完了する。

5.2 階層型セグメントルーティングの実装

本節では第 5.1 節で行った設計を元に階層型セグメントルーティングの実装を行う。本構成では SR-MPLS を選択し、データプレーンには Linux Kernel・iproute2 に実装されている MPLS を、またコントロールプレーンは FRRouting に実装されている OSPF-SR を利用する。

本研究では AS 連携時の管理コストの利点と実装負荷を考慮し、階層型セグメントルーティングのみでサービスチェイニングが実現可能であり、Node SID と Adjacency SID で実装可能である展開モデルを採用した。また、ポリシ設定手法も同様に AS 連携を考慮し、収集モデルを選択する。

5.2.1 階層型 SR PCE の実装

階層型 SR PCE の上位 SR PCE は既存の SR PCE を利用するため、まず既存の SR PCE の実装を行う。

既存の SR PCE の実装を図 5.3 に示す。実装と同様に、図の左型に TED 更新処理を、図の右側にセグメントリスト構築処理を示した。TED 更新処理では、まず SR PCC はリンクステート取得処理を実行し、vtysh 経由で FRRouting の OSPF で用いられる LSDB を参照し、リンクステートを取得する。vtysh の結果は文字列型として得られるため、構文解析を行い配列型に変換する。その後、取得したリンクステートを SR PCE へ送信する。本実装では、BGP-LS ではなく TCP ソ

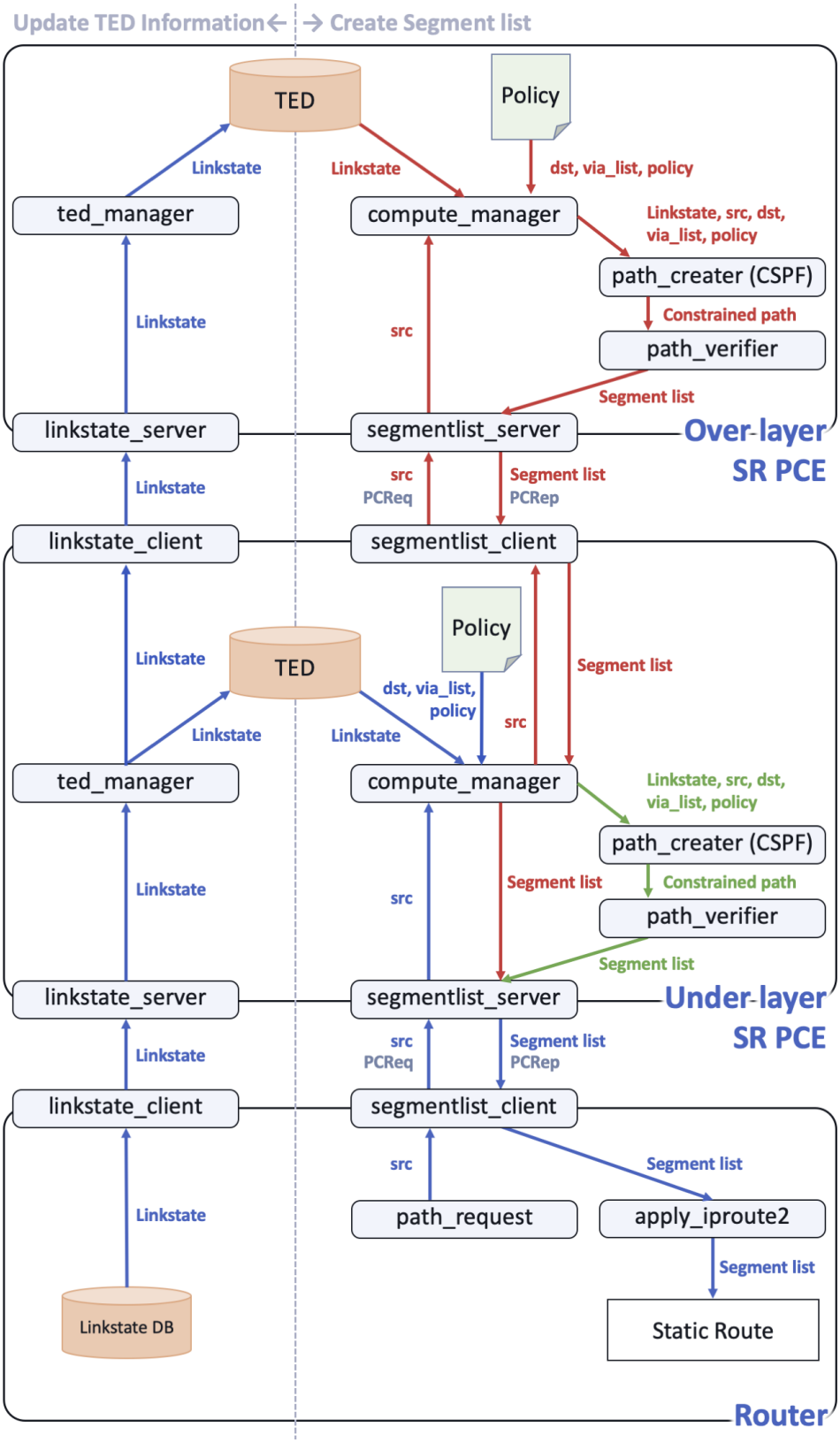

図 5.2: 階層型 SR PCE の設計

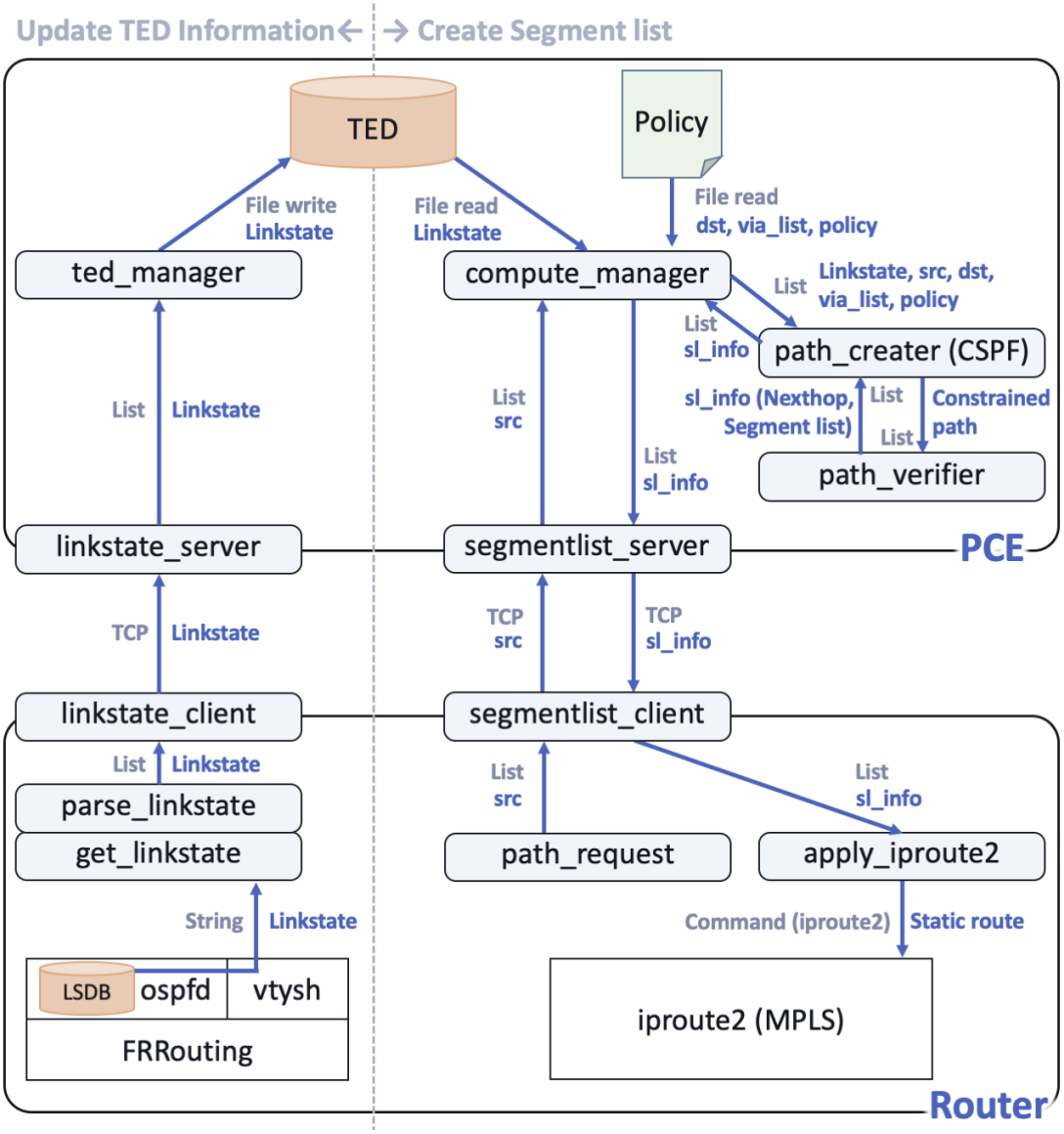

図 5.3: 既存 SR PCE の実装

ケットを採用した。SR PCE が受け取ったリンクステートを TED に格納することで TED 更新処理が完了する。

セグメントリスト構築処理では、まず SR PCC が PCReq を発行する。本実装では PCReq に PCEP ではなく TCP ソケットを採用した。PCReq を受信した SR PCE では TED とポリシーを参照し、リンクステートとポリシー、宛先の情報を取得する。取得した情報を元に CSPF 計算を行うことで制約付き経路を構成し、それを満たすようにセグメントリストを構築する。この際、セグメントリストに加え、隣接ノードのアドレス情報を出力する。その後 PCReq として、TCP ソケットを利用しセグメントリストと隣接ノードのアドレスを SR PCC へと送信する。SR PCC が受け取ったセグメントリストと隣接ノードの情報を利用し、iproute2 のコマンドによりスタティックルートの設定を行うことでセグメントリスト構築処理が完了する。

階層型 SR PCE の実装を図 5.4 に示す。設計と同様に階層型 SR PCE の上位 SR PCE は既存の SR PCE を用い、下位 SR PCE は既存の SR PCE に新たな処理を追加し作成する。

TED 更新処理では、下位 SR PCE は SR PCC からリンクステートを受け取り、TED に格納すると同時に上位 SR PCE へ送信を行う。セグメントリスト構築処理でも設計と同様に、下位 SR PCE から PCReq を受け取った際、上位 SR PCE への転送処理と計算処理を行う。計算処理では、TED とポリシーを参照し、送信元に対応する全ての送信先へのセグメントリストを構築し、隣接ノード情報と共に SR PCE へと送信する。また、上位 SR PCE から受信したセグメントリスト・隣接ノード情報も同様に SR PCE へと送信を行う。

SR PCC と上位・下位 SR PCE を含めた各 SR PCE の処理を表したシーケンス図を図 5.5 に示す。凡例の通りにそれぞれの処理を記号で表す。リンクステート更新処理では FRRouting を A、SR PCC を C、下位 SR PCE を E、上位 SR PCE を G と表現する。セグメントリスト更新処理では iproute2 を B、SR PCC を D、下位 SR PCE を F、上位 SR PCE を H と表現する。また、下位 SR PCE の TED を α、上位 SR PCE の TED を β と表す。

図 5.5 右上のリンクステート更新処理では線形に処理が行われる。FRRouting である A から SR PCC である C がリンクステートを受け取り、下位 SR PCE である E へと送信する。下位 SR PCE は TED である α への格納と上位 SR PCE である G への送信を並列に行い、上位 SR PCE が受信したリンクステートを TED である β に格納することで処理が完了する。

図 5.5 右下のセグメントリスト構築処理では、分岐処理により 2 つの返答が行われる。まず SR PCC である D が下位 SR PCE である F へ PCReq を送信する。下位 SR PCE は上位 SR PCE である H への PCReq の送信と、経路計算のため TED である α からのリンクステート取得を並列で行う。その後下位 SR PCE はセグメントリストを構築し、隣接ノード情報と共に下位 SR PCC へと送信する。上位 SR PCE は PCReq を受信した後、TED である β からリンクステートを取得し、経路

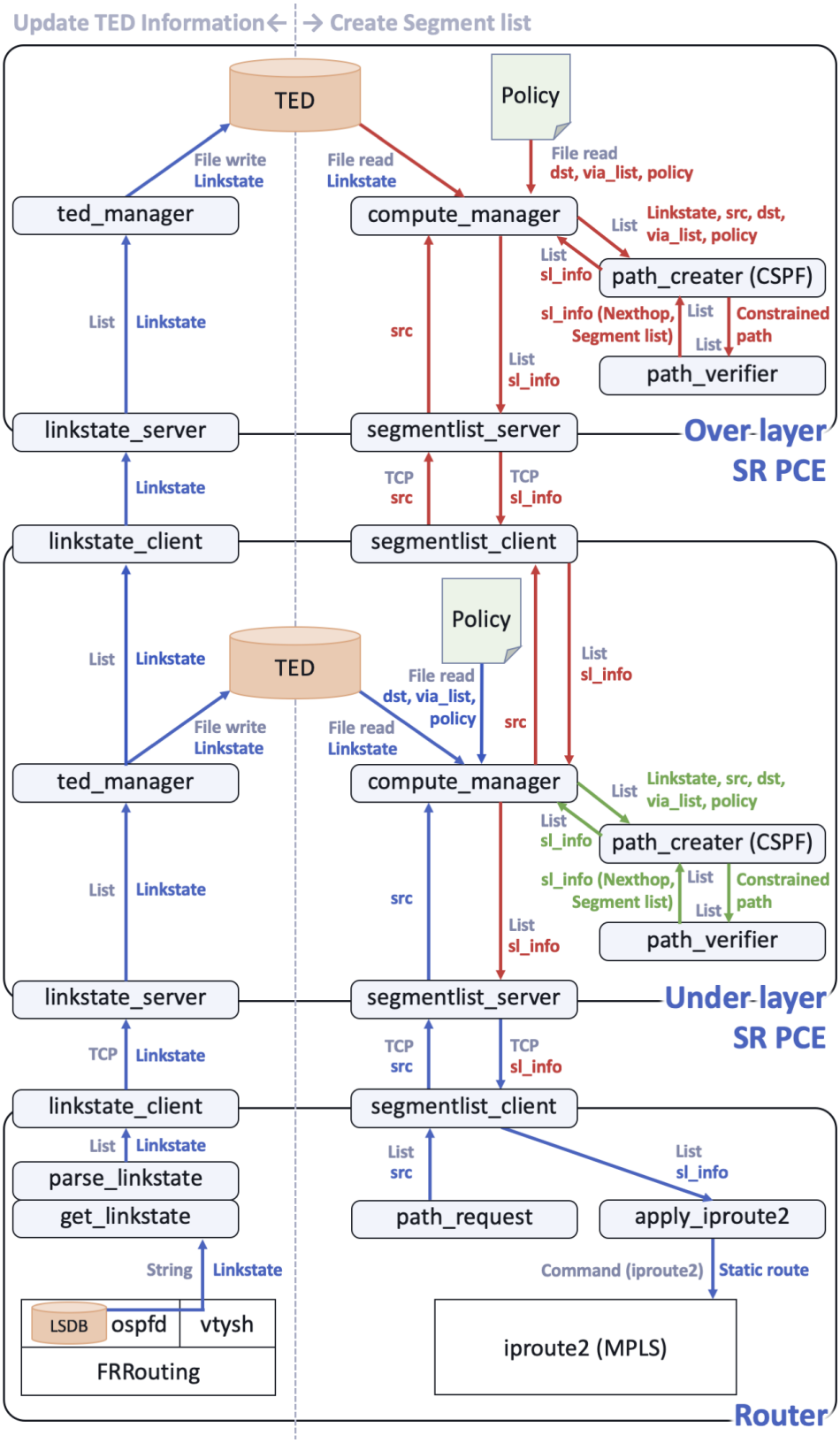

図 5.4: 階層型 SR PCE の実装

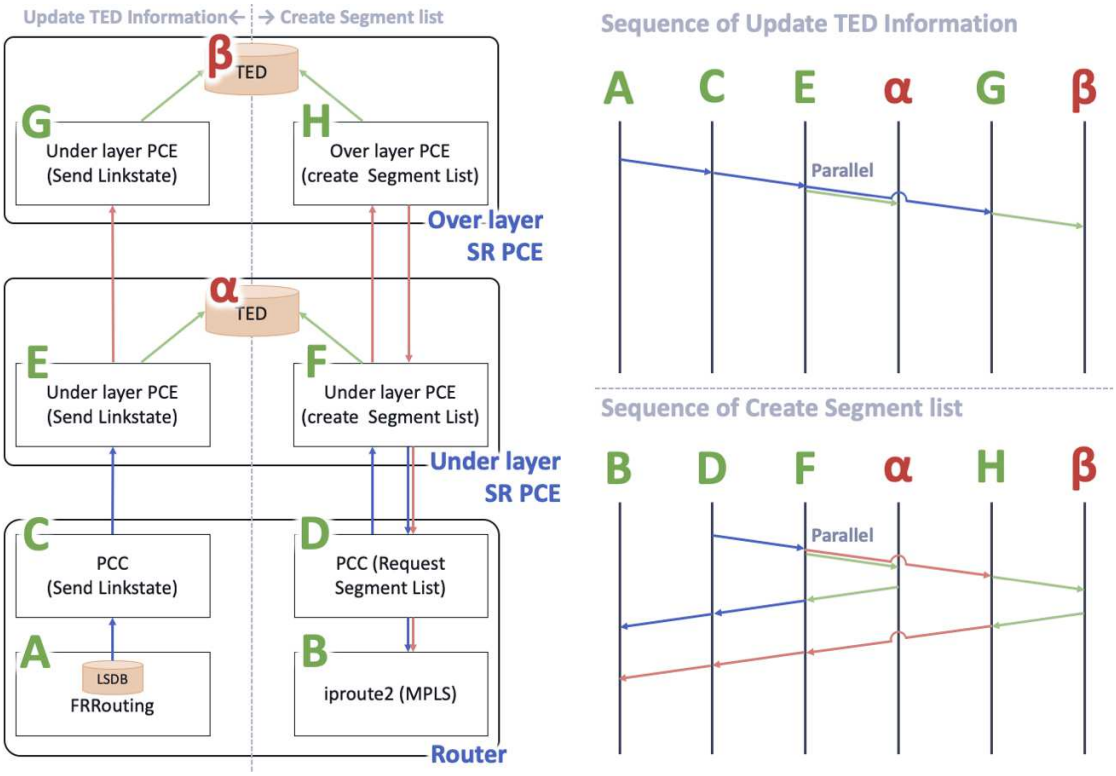

図 5.5: 階層型 SR PCE のシーケンス図

計算を行いセグメントリストを構築する。構築したセグメントリストと隣接ノードの情報を下位 SR PCE へと送信する。下位 SR PCE は上位 SR PCE から情報を受け取り、SR PCC へと送信する。SR PCC は下位 SR PCE からセグメントリストと隣接ノードの情報を受け取った際に iproute2 を実行しスタティックルートを作成する。

5.2.2 サブドメイン間のリンクステート取得

本構成では、コントロールプレーンに FRRouting の OSPF-SR を利用する。OSPF は AS 内で利用する IGP であるため、図 5.6 に示したように、AS 間のリンクステート管理や Adjacency Segment の付与を行うことができない。この課題の解決のため、複数の手法が考えられる。実装を行う上でそれらの手法の検討を行う。

図 5.7 に 4 種の解決手法を示す。1 つ目は、BGP 等の External Gateway Protocols（EGPs）でリンクステートの取得・Adjacency SID の付与を行う手法である。AS 連携に階層型セグメントルーティングを用いる際、AS 間のリンクには BGP 等の EGPs が設定されると予想される。BGP では対向ノードとの隣接関係を構築するため Adjacency SID の発行を自然に行うことが可能となり、またリンクの予約可能帯域等の QoS 情報なども取り決めることが可能となる。この手法の課題として、

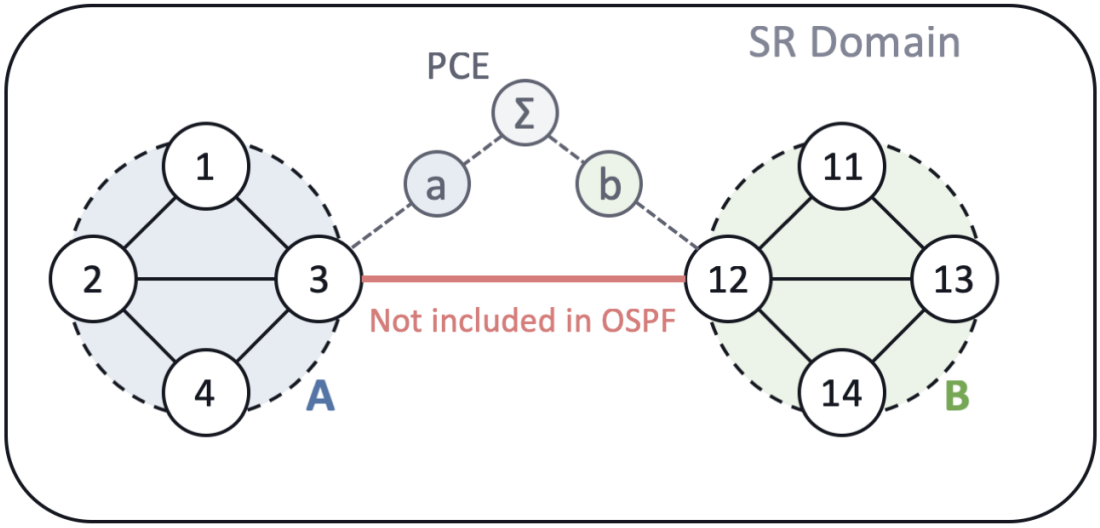

図 5.6: AS 間のリンクステート取得に対する課題

IGP と EGPs の双方にセグメントルーティングの設定を行う必要があり、管理コストや運用コストが上昇する。

2 つ目は、サブドメイン間のインターフェースをパッシブインターフェースに設定することで、IGP によりリンクステートを取得する手法である。パッシブインターフェースとはトラフィック量の削減や想定外の経路広告を受け取ることを防ぐために、ルーティングプロトコルの管理対象に含めるが該当インターフェースからの経路広告をせず、また経路広告を受け取らない設定である。OSPF 等の IGP では、管理のため AS 間のリンク等の OSPF の経路広告を行わないインターフェースにもパッシブインターフェースを指定することが考えられる。パッシブインターフェースに設定することにより、OSPF-TE の予約可能帯域等の QoS 情報を得ることが可能となる。しかしこの手法と課題として、隣接関係を作成しないため Adjacency SID を別手法で設定する必要がある点が挙げられる。

3 つ目は、各サブドメインのエッジルータで IGP のインスタンスを複数起動し、所属するサブドメインの IGP に加え、隣接するサブドメインの IGP 参加させる手法である。この手法ではサブドメインのエッジルータ同士が IGP の隣接関係を構築するため、経路制御のための QoS 情報と隣接関係による Adjacency SID の割り振りの両方を IGP のみで実現できる。しかし階層型セグメントルーティングの利点であったサブドメイン毎の設定分割が不可能となるデメリットが生じるため、本研究の提案には適さない手法である。

BGP が動作する環境では 1 つ目の手法が適しており、IGP のみでセグメントルーティングを動作させる場合は 2 つ目の手法が適している。本研究では、IGP のみで実現可能である利点から、2 つ目の手法であるパッシブインターフェースを選択し、Adjacency SID は手動で割り当てることとする。

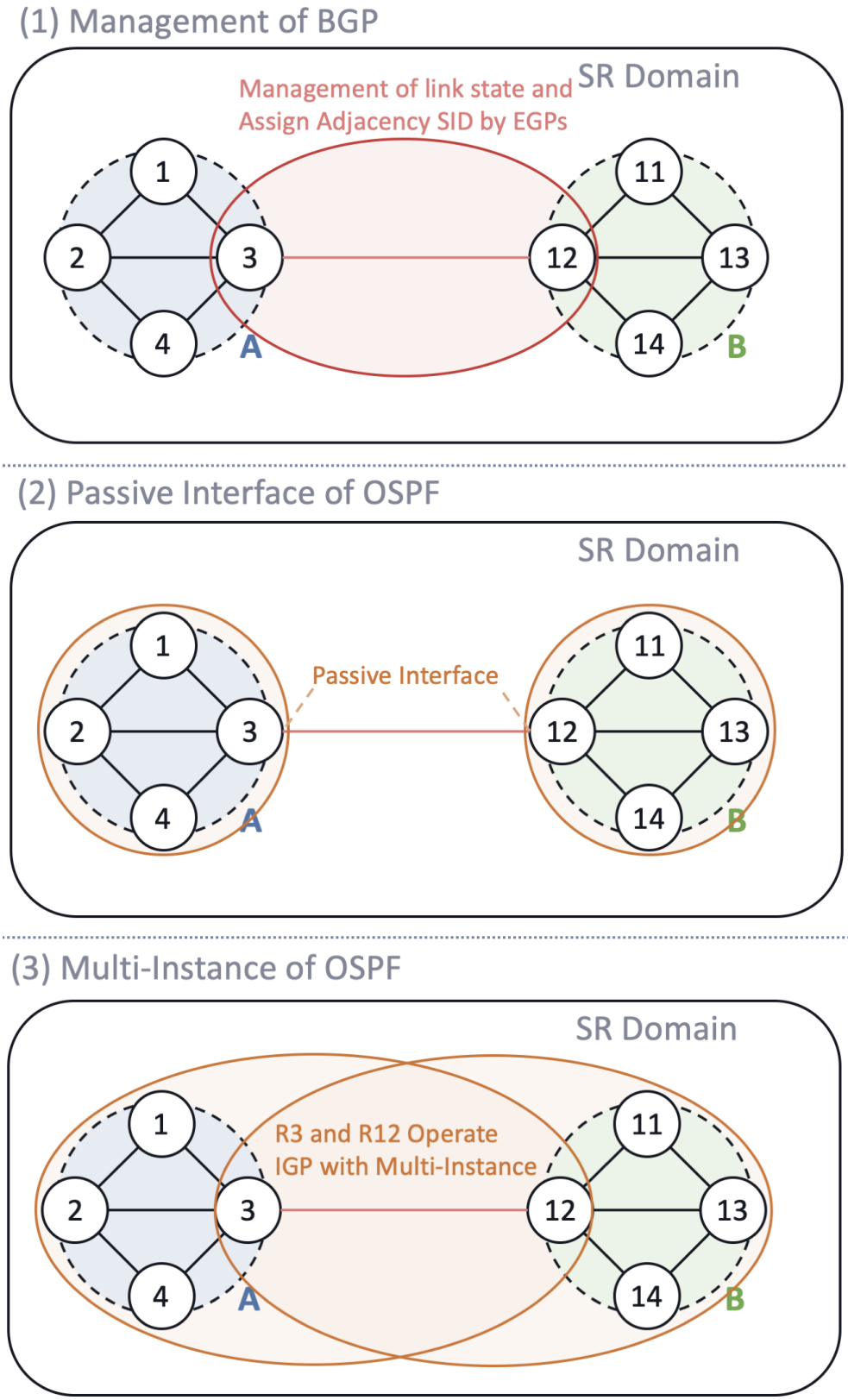

図 5.7: AS 間のリンクステート取得手法

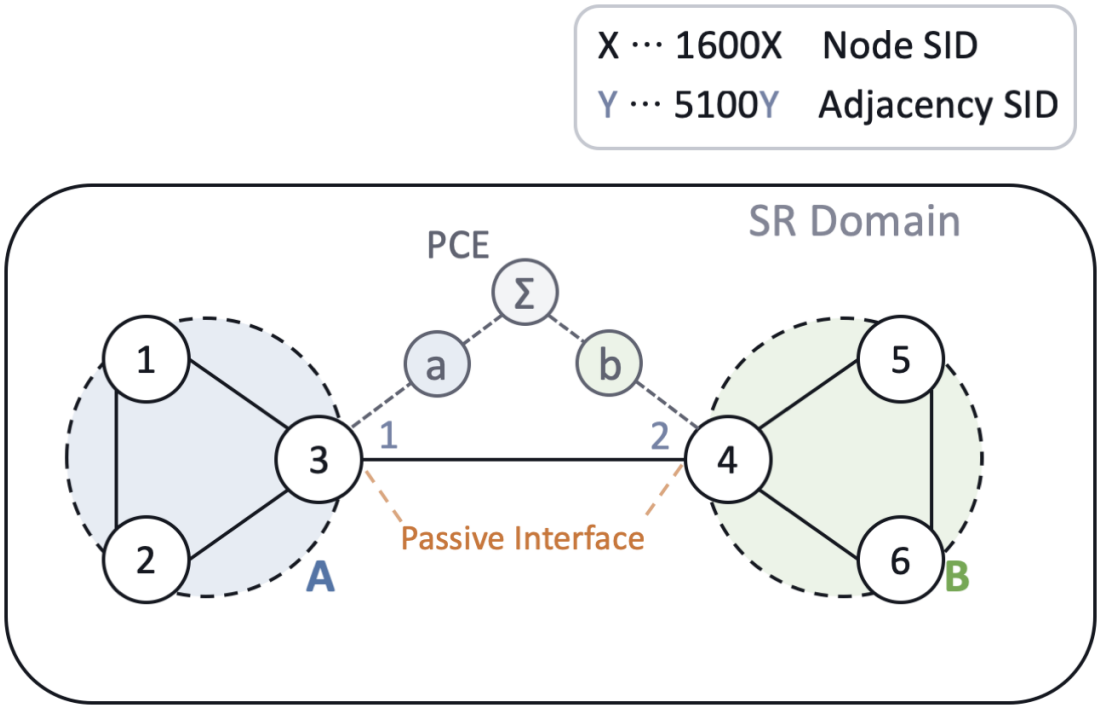

図 5.8: 展開モデル動作用トポロジ

5.2.3 階層型セグメントルーティングの動作

本項では階層型 SR PCE に実装した展開モデルの動作を示す。図 5.8 のネットワークを作成し、実装した階層型セグメントルーティングを動作させる。

例では 2 つのサブドメインを作成した。サブドメイン A にはノード 1・2・3 が、サブドメイン B にはノード 4・5・6 がそれぞれ含まれており、ノード 3 とノード 4 がサブドメイン間のリンクを有している。このトポロジにおいて、ノード 1 からノード 6 への通信に以下の表 5.1 に示したポリシー例（1）を設定する。

経由ノードは表記した順番に経由するものとする。ポリシー例（1）が適用された場合、図 5.8 に示したようにノード 1 から 6 への経路は 1 - 2 - 3 - 4 - 5 - 6 の順となる。

表 5.1: ポリシー例（1）

項目名	値
送信元	ノード 1
宛先	ノード 6
経由ノード	ノード 2・ノード 5
帯域制限	無し
回避ノード	無し

```
watal@r1:~/python-pcc$ python python_pcc.py
[Segment list] Finished sending request
[Segment list] Finished receiving segmentlist infomation
watal@r1:~/python-pcc$ ip r | grep 192.168.0.6
192.168.0.6 encap mpls 16003/51001/16005/16006 via 172.16.0.2 dev ens4
```

図 5.9: ポリシー例（1）適用結果

表 5.2: ポリシー例（2）

項目名	値
送信元	ノード 1
宛先	ノード 6
経由ノード	ノード 4・ノード 2・ノード 5
帯域制限	無し
回避ノード	無し

ポリシー 1 に基づき展開モデルで経路計算を行うと、隣接ノードであるノード 2 へ、セグメントリスト<16003, 51001, 16005, 16006>が設定される。図 5.9 に実行結果を示す。実行結果からセグメントリストが正しく構築され、ノード 2 と隣接するインターフェースである ens4 が指定されていることがわかる。

次に、このトポロジにおいてノード 1 からノード 6 への通信に以下の表 5.2 に示したポリシー例（2）を設定する。ポリシー例（2）が適用された場合、隣接ノードであるノード 3 へ、セグメントリスト<51001, 51002, 16002, 16003, 51001, 16005, 16006>が設定される。図 5.10 に実行結果を示す。実行結果からセグメントリストが正しく構築され、ノード 3 と隣接するインターフェースである ens5 が指定されていることがわかる。

```
watal@r1:~/python-pcc$ python python_pcc.py
[Segment list] Finished sending request
[Segment list] Finished receiving segmentlist infomation
watal@r1:~/python-pcc$ ip r | grep 192.168.0.6
192.168.0.6 encap mpls 51001/51002/16002/16003/51001/16005/16006 via 172.16.0.6
     dev ens5
```

図 5.10: ポリシー例（2）適用結果

第6章　評価

本章では、提案手法の評価実験・性能評価・既存手法との比較について述べる。

6.1　階層型セグメントルーティングの効果

本研究の提案手法である階層型セグメントルーティングにより、セグメントルーティングにおけるサブドメインの作成とサブドメインを越えた経路制御を実現した。本節では研究の目的である AS 連携による評価を行う。

6.1.1　AS ごとの分離

階層型セグメントルーティングの AS 連携について、複数 AS が存在する環境への適用例を用い評価を行う。

図 6.1 に複数の AS が存在する環境を示す。この環境には ISP-S・ISP-D という 2 つの AS が存在し、階層型セグメントルーティングにより連携を行う。ISP-S には host1 とノード 1・2・3 が、ISP-D には host2 とノード 4・5・6 が存在する。また、ノード 2 とノード 5 ではそれぞれ VNF が提供されている。このトポロジにおいて host1 から host2 へと通信を行う際、最短経路である 1 - 3 - 4 - 6 の経路ではなく、ノード 2、ノード 5 を経由するサービスチェインを提供するため、1 - 2 - 3 - 4 - 5 - 6 という経路を構成する。

この例における ISP-S 内の設定を表 6.1 に、ISP-D 内の設定を表 6.2 に示す。AS ごとの ID 空間の分離を示すため、本環境では ISP-S・ISP-D の双方で Node SID を 16001 から 16003 の範囲で割り振りを行い、また各ノードのルータ ID も 192.168.0.1・192.168.0.2・192.168.0.3 を使用した。表 6.1・表 6.2 の例からわかる通り、異なる AS である ISP-S と ISP-D の OSPF 設定は各自の AS のみを対象としており、他の AS や AS 間のリンクである 172.16.0.12/30 の経路はどちらの OSPF にも広告されていないため AS ごとの設定独立が実現されている。また Node SID の空間も分離している。

本設定適用後、ISP-S・ISP-D 内ノードの SID テーブルの確認を行う。図 6.2 に ISP-S のルータであるノード 1 の SID テーブルを、また図 6.3 に ISP-D のルータであるノード 4 の SID テーブルを示す。

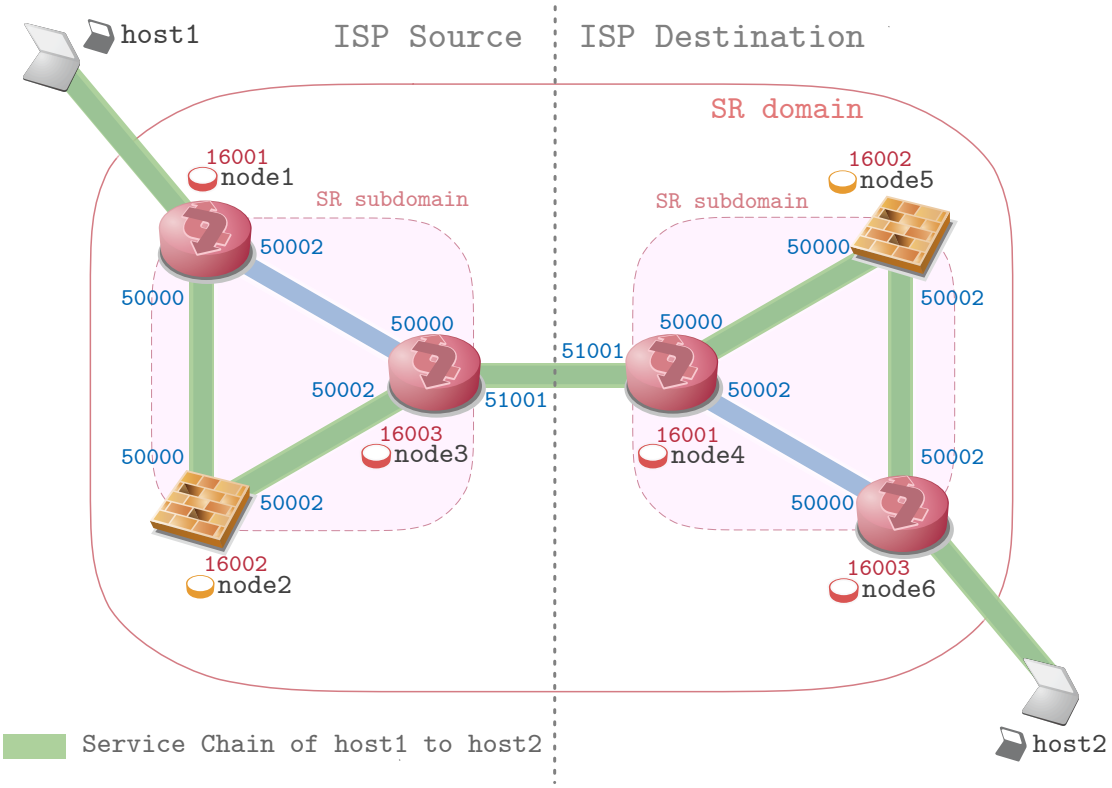

図 6.1: AS 連携評価用トポロジ

表 6.1: ISP-S 内の設定

設定項目		ノード 1	ノード 2	ノード 3
インターフェース		172.16.0.1/30	172.16.0.2/30	172.16.0.6/30
		172.16.0.5/30	172.16.0.9/30	172.16.0.10/30
				172.16.0.13/30
OSPF	Area	0	0	0
	Router-ID	192.168.0.1	192.168.0.2	192.168.0.3
	Advertising NW	172.16.0.0/30	172.16.0.0/30	172.16.0.4/30
		172.16.0.4/30	172.16.0.8/30	172.16.0.8/30
		192.168.0.1/32	192.168.0.2/32	192.168.0.3/32
	Node SID	16001	16002	16003
	SRGB	16000-20000	16000-20000	16000-20000

表 6.2: ISP-D 内の設定

設定項目		ノード 4	ノード 5	ノード 6
インターフェース		**172.16.0.14/30**	172.16.0.18/30	172.16.0.22/30
		172.16.0.17/30	172.16.0.25/30	172.16.0.26/30
		172.16.0.21/30		
OSPF	Area	0	0	0
	Router-ID	192.168.0.1	192.168.0.2	192.168.0.3
	Advertising NW	172.16.0.l6/30	172.16.0.16/30	172.16.0.20/30
		172.16.0.20/30	172.16.0.24/30	172.16.0.24/30
		192.168.0.1/32	192.168.0.2/32	192.168.0.3/32
	Node SID	16001	16002	16003
	SRGB	16000-20000	16000-20000	16000-20000

ノード 1 の SID テーブルにはノード 2・ノード 3 の Node SID である 16002・16003 と、各 Adjacency SID が格納されている。FRRouting では各隣接関係にバックアップを含めた 2 つの Adjacency SID を付与するため、ここではノード 2 への隣接関係を示す 50001・50002、ノード 3 への隣接関係を示す 50003・50004 の Adjacency SID が格納されている。これにより、ノード 1 の保有する情報が ISP-S 内部に限定されていることが確認できる。

ノード 4 の SID テーブルにはノード 5・ノード 6 の Node SID である 16002・16003 と、ノード 3 への隣接関係を示す 51001・51002、ノード 5 への隣接関係を示す 50000・50001、ノード 5 への隣接関係を示す 50002・50003 の Adjaency SID が格納されている。本実験ではサブドメイン間はパッシブインターフェースとして設定し、Adjacency SID は iproute2 を利用し設定している。そのため、Adjacency SID のうち OSPF により生成された行には proto ospf と表示されており、サブド

```
watal@r1:~$ ip -M route
16002 via inet 172.16.0.2 dev ens4 proto ospf
16003 via inet 172.16.0.6 dev ens5 proto ospf
50000 via inet 172.16.0.6 dev ens5 proto ospf
50001 via inet 172.16.0.6 dev ens5 proto ospf
50002 via inet 172.16.0.2 dev ens4 proto ospf
50003 via inet 172.16.0.2 dev ens4 proto ospf
```

図 6.2: ノード 1 の SID テーブル

```
watal@r4:~$ ip -M route
16002 via inet 172.16.0.18 dev ens5 proto ospf
16003 via inet 172.16.0.22 dev ens6 proto ospf
50000 via inet 172.16.0.18 dev ens5 proto ospf
50001 via inet 172.16.0.18 dev ens5 proto ospf
50002 via inet 172.16.0.22 dev ens6 proto ospf
50003 via inet 172.16.0.22 dev ens6 proto ospf
51001 via inet 172.16.0.13 dev ens4
51002 via inet 172.16.0.13 dev ens4
```

図 6.3: ノード 4 の SID テーブル

メイン間の Adjacency SID である 51001・51002 には proto ospf の表示は無い。図 6.3 により、ノード 4 の保有する情報が ISP-D 内部と、自らの Adjacency SID のみに限定されていることが確認できる。

これらの例より、各ノードの SID テーブル・ID 空間がサブドメインである AS ごとに分離されていることが確認できた。例では各ノードはサブドメイン内の Node SID と Adjacency SID のみを保持しているため、転送は展開モデルを用い実現する。

ノード 1 からノード 6 への通信に適用するポリシーを図 6.3 に示す。このポリシーを適用した場合、それぞれのセグメントリストは図 6.4、図 6.5 のようになる。ノード 1 からノード 6 へのセグメントリストでは、最短経路である 1 - 3 - 4 - 6 ではなく、ノード 2・ノード 5 を通る 1 - 2 - 3 - 4 - 5 - 6 の経路が実現されていることが確認できる。同様に、ノード 6 からノード 1 へのセグメントリストも最短経路である 6 - 4 - 3 - 1 ではなく、ノード 5・ノード 2 を通る 6 - 5 - 4 - 3 - 2 - 1 の経路が実現されている。例より、各ノードの設定・経路広告範囲の分離と情報削減

表 6.3: ノード 1・ノード 6 間のポリシー

宛先	ノード 6	ノード 1
経由ノード	ノード 2・ノード 5	ノード 5・ノード 2
帯域制限	無し	無し
回避ノード	無し	無し

```
watal@r1:~/python-pcc$ python3 python_pcc.py
[Segment list] Finished sending request
[Segment list] Finished receiving segmentlist infomation
shinoda-lab@r1:~/python-pcc$ ip r | grep 192.168.0.6
192.168.0.6 encap mpls 16003/51001/16002/16003 via 172.16.0.2 dev ens4
```

図 6.4: ノード 1 からノード 6 へのセグメントリスト

```
shinoda-lab@r6:~/python-pcc$ python3 python_pcc.py
[Segment list] Finished sending request
[Segment list] Finished receiving segmentlist infomation
shinoda-lab@r6:~/python-pcc$ ip r | grep 192.168.0.1
192.168.0.1 encap mpls 16001/51001/16002/16001 via 172.16.0.25 dev ens5
```

図 6.5: ノード 6 からノード 1 へのセグメントリスト

が行われ、その上で経路制御が実現されたことを示した。これらの検証より、本研究の課題である AS 連携における課題の解決を確認した。

6.1.2 AS を越えたサービスチェイニングへの適用

階層型セグメントルーティングの適用例として AS を越えたサービスチェイニングを実現し、その評価を行う。図 6.6 に、第 2.3.1 項で触れた複数の AS に VNF が分散したトポロジを示す。本検証のシナリオとして、ISP-S・ISP-D・パブリッククラウドの 3 つの AS が業務連携し、一連のサービスチェインを顧客に提供することを考える。ここでサービスチェインの構成要素は、3 つの AS が提供する全ての機能から顧客が自由に選択することを想定する。図 6.6 では、パブリッククラウドが DPI を提供するノード c-R1 とファイアウォールを提供するノード c-R3 を、また、ISP-S がファイアウォールを提供するノード s-R4 をそれぞれ運用している。

このシナリオに基づき、階層型セグメントルーティングを適用することで各 AS の分割を行いながらサービスチェイニングを実現する。図左のパブリッククラウドは 1 つのサブドメインとして構成されている。また、ISP-S・ISP-D ではさらに内部でサブドメインが分離を行なっている。このトポロジでは、顧客のホームルータもセグメントルーティングに対応し、SR ドメインに参加することを想定する。また、説明を単純化するため各 VNF も SR ドメインに参加させる。各顧客の情報を分離し、顧客同士の情報や ISP のコアネットワークの情報がお互いに公開されることを防ぐ目的と、顧客のホームルータの要求性能を下げるため、ホームルータとなる各 host を収容するノード s-R5・s-R6 や、Webserver を収容するノード d-R1 は 1 つのサブドメインとして構成する。また、図 6.6 には host1・host2 それぞれから Webserver までの最短経路を示している。このトポロジにおいて、host1、host2

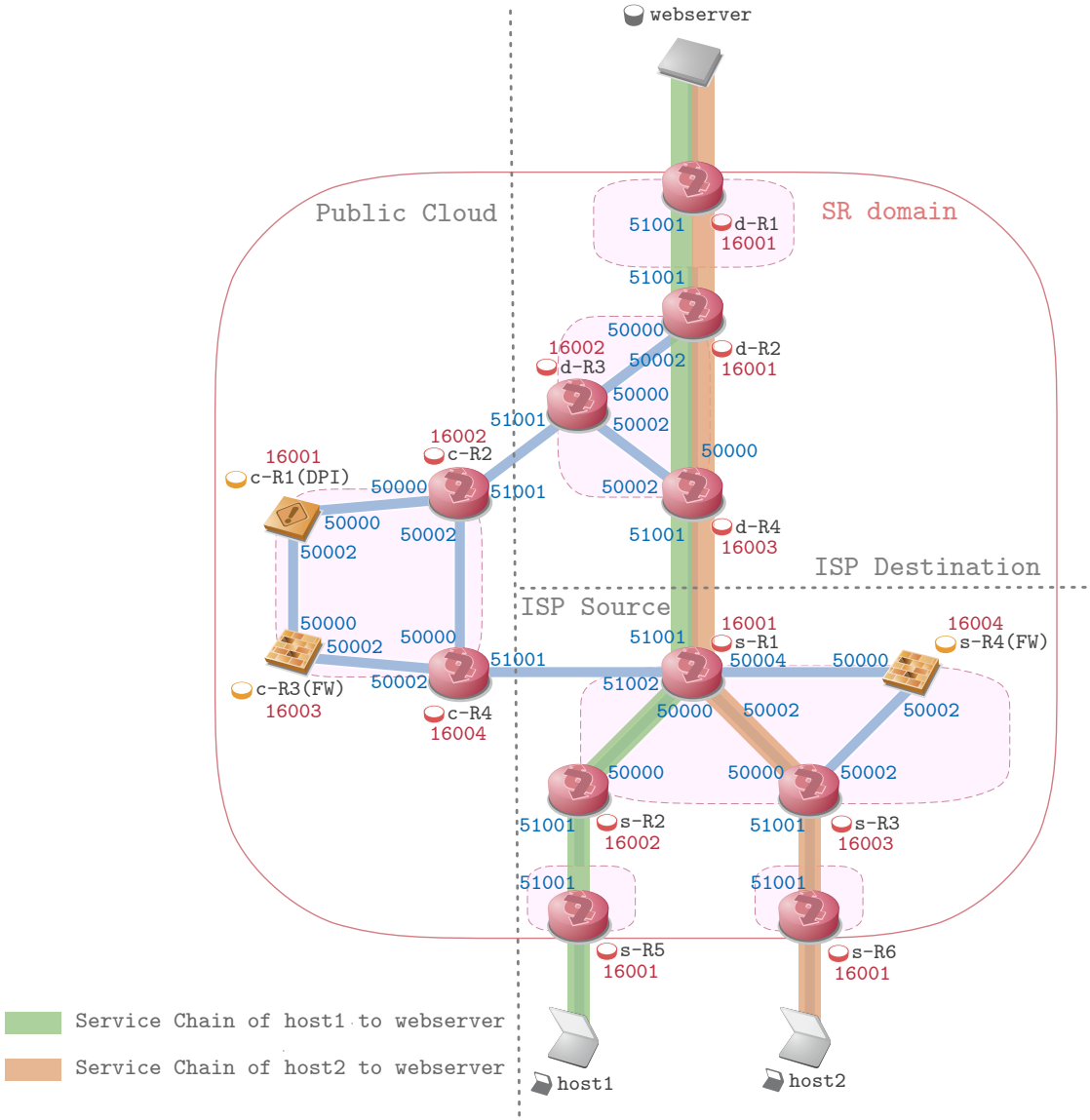

図 6.6: 複数の AS に VNF が分散したトポロジ

表 6.4: 各ホストのサービスチェイン用ポリシー

送信元	host1	host2
宛先	Webserver	Webserver
経由ノード	c-R3 (Firewall)・c-R1 (DPI)	s-R4 (Firewall)・c-R1 (DPI)
帯域制限	無し	無し
回避ノード	無し	無し

から Webserver へのサービスチェイニングを実現する。また、それぞれのサービスチェインは表 6.4 に示したポリシーに基づき構成する。host1 はファイアウォール・DPI の双方をパブリッククラウドの提供する機能を選択するため、c-R3 と c-R1 が経由ノードに指定されている。それに対し host2 ではファイアウォールは ISP-S が提供する機能を選択し、DPI はパブリッククラウドの提供する機能を選択するため、s-R4 と c-R1 が経由ノードとして指定されている。

図 6.7 にサービスチェイニング実現時の経路を示す。例より、サービスチェインの提供により host1、host2 から Webserver への経路が変更され、要求通りに VNF を経由していることが確認できる。host1 から Webserver へのサービスチェインを実現するセグメントリストを図 6.8、図 6.9 に、host2 から Webserver へのサービスチェインを実現するセグメントリストを図 6.10、図 6.11 に示す。

本検証では AS ごとに ID 空間を分割しており、各 AS で同じ ID を別ノードに割り振っているため、セグメントリストにおいても同じ SID が別のノードを示すために用いられる。例えば、図 6.8 では、セグメントリストの先頭要素である 16001 は s-R1 を、4 番目の要素である 16001 は c-R1 を示している。例より、図 6.8、図 6.9 では c-R3 と c-R1 が、図 6.10、図 6.11 では s-R4 と c-R1 がセグメントリスト中に指定されており、表 6.4 に示したポリシーに従いセグメントリストが構築されていることが確認できる。これらの検証結果より、階層型セグメントルーティングを用いた AS 連携によるサービスチェイニングの実現を確認した。

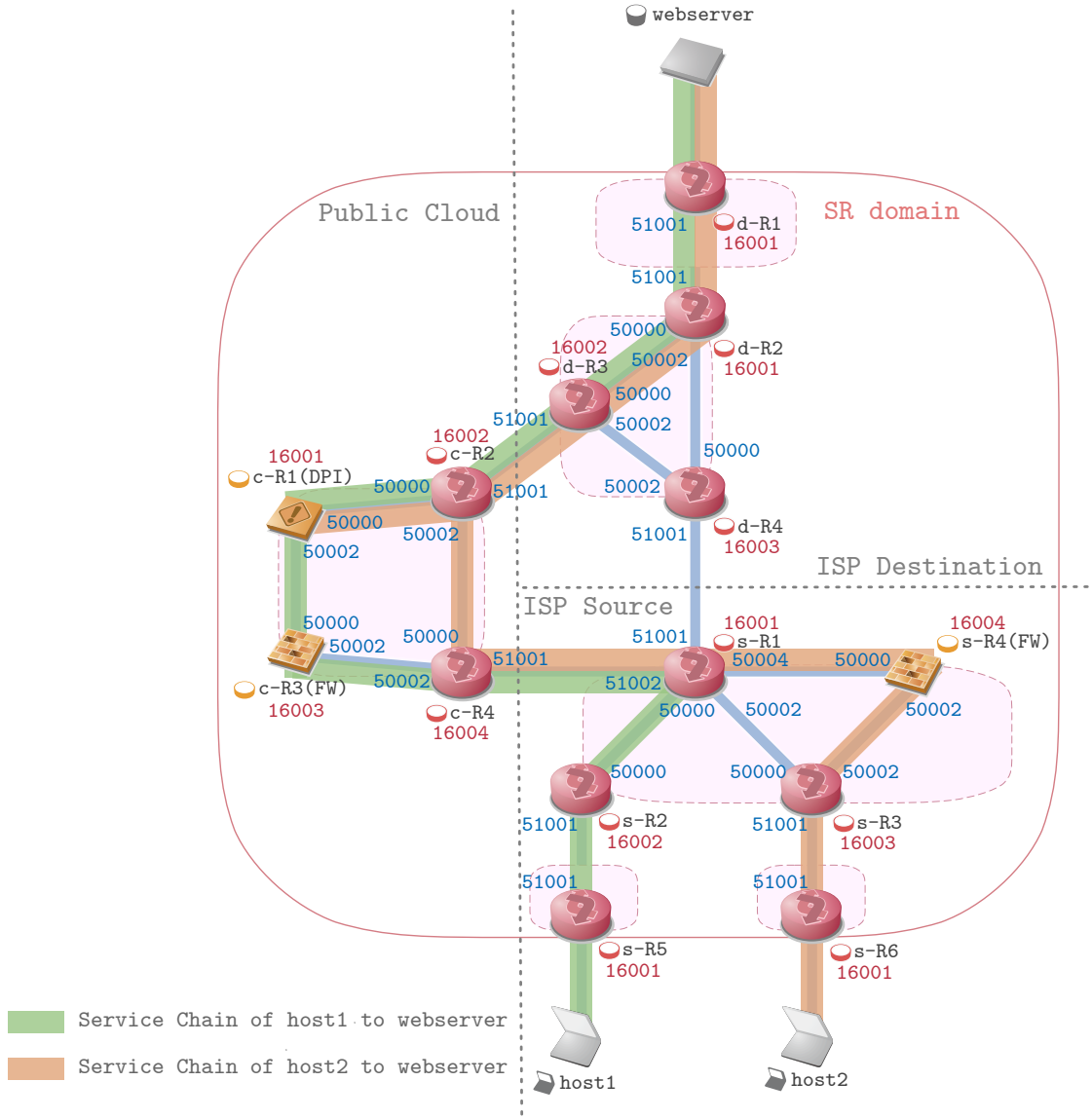

図 6.7: サービスチェイニング実現時の経路

```
watal@isp-s-r5:~$ ip r | grep 172.16.0.0/30
172.16.0.0/30 encap mpls 16001/51002/16003/16001/16002/51001/16001/51001 via
    172.16.1.17 dev ens4
```

図 6.8: host1 から Webserver へのセグメントリスト

```
watal@isp-d-r1:~$ ip r | grep 172.16.1.24/30
172.16.1.24/30 encap mpls 16002/51001/16001/16003/16004/51001/16002/51001 via
    172.16.0.6 dev ens5
```

図 6.9: Webserver から host1 へのセグメントリスト

```
watal@isp-s-r6:~$ ip r | grep 172.16.0.0/30
172.16.0.0/30 encap mpls 16004/16001/51002/16001/16002/51001/16001/51001 via
    172.16.1.21 dev ens4
```

図 6.10: host2 から Webserver へのセグメントリスト

```
watal@isp-d-r1:~$ ip r | grep 172.16.1.28/30
172.16.1.28/30 encap mpls 16002/51001/16001/16004/51001/16004/16003/51001 via
    172.16.0.6 dev ens5
```

図 6.11: Webserver から host2 へのセグメントリスト

表 6.5: 計測環境

Host	Model	PowerEdge R620
	CPU	Intel(R) Xeon(R) CPU E5-2650 0 @ 2.00GHz
	Memory	192 GBytes
	OS	Ubuntu 18.10 Server
	Virtualizer	QEMU/KVM
Guest VM	Virtual CPU	1
	Virtual Memory	512 MBytes
	OS	Ubuntu 16.04 Server

6.2 階層型SR PCE

本節では階層型SR PCEについて、AS連携時に必要となる規模追従性を評価する。性能評価を行うにあたり、使用した環境を表6.5に示す。

ASのモデルとして、本研究ではInternet2のIPネットワーク[21]を用いる。図6.12にトポロジを示す。

図6.12のトポロジよりRouter Nodeを抜き出すと、図6.13上段に示したトポロジとなる。このトポロジを1つのサブドメインとして、図6.13中段・下段に示したように複数のサブドメインが連携した場合のノード数増加とセグメントリスト構築時間の関係を計測し評価する。ここでは実装を行なった展開モデルの階層型SR PCEを用いる。また、比較対象としてサブドメイン分割を行わない既存のSR PCEを用いる。図6.14に計測結果を示す。結果より、階層型SR PCE・既存のSR PCEの双方において、ノード数の増加に伴いセグメントリスト構築時間が増加していることが確認できる。優先度付きキューを用いた場合のDijkstraアルゴリズムの計算量は$\mathcal{O}(n \log n)$となる。展開モデルではドメイン全体のDijkstra計算に加え、経路検証のため経路をサブドメインごとに分割し複数回の経路計算を行うため、既存のSR PCEよりも計算量が増加する。Dijkstraアルゴリズムの対象となるトポロジをm個に分割し、m回に分けて計算を行う際の計算量を以下に示す。

$$\mathcal{O}\left(m \times \frac{n}{m} \log \frac{n}{m}\right) = \mathcal{O}\left(n \log \frac{n}{m}\right)$$

また、結果において階層型SR PCEと既存のSR PCEの経過時間が近接している箇所が存在するが、これはサブドメイン分割の結果、トポロジの形によっては経路に含まれないサブドメインが生じ、経路計算に用いるDijkstraアルゴリズムの計算範囲が減少するためである。

本実験における36ノードの場合、トポロジは図6.13下のように接続されている。このため、ノードSが存在するサブドメインAとノードGが存在するサブドメインDではDijkstraアルゴリズムの計算が行われるが、その他のサブドメインは計算対象とならないため、計算量の削減が生じている。今回の実装ではサブド

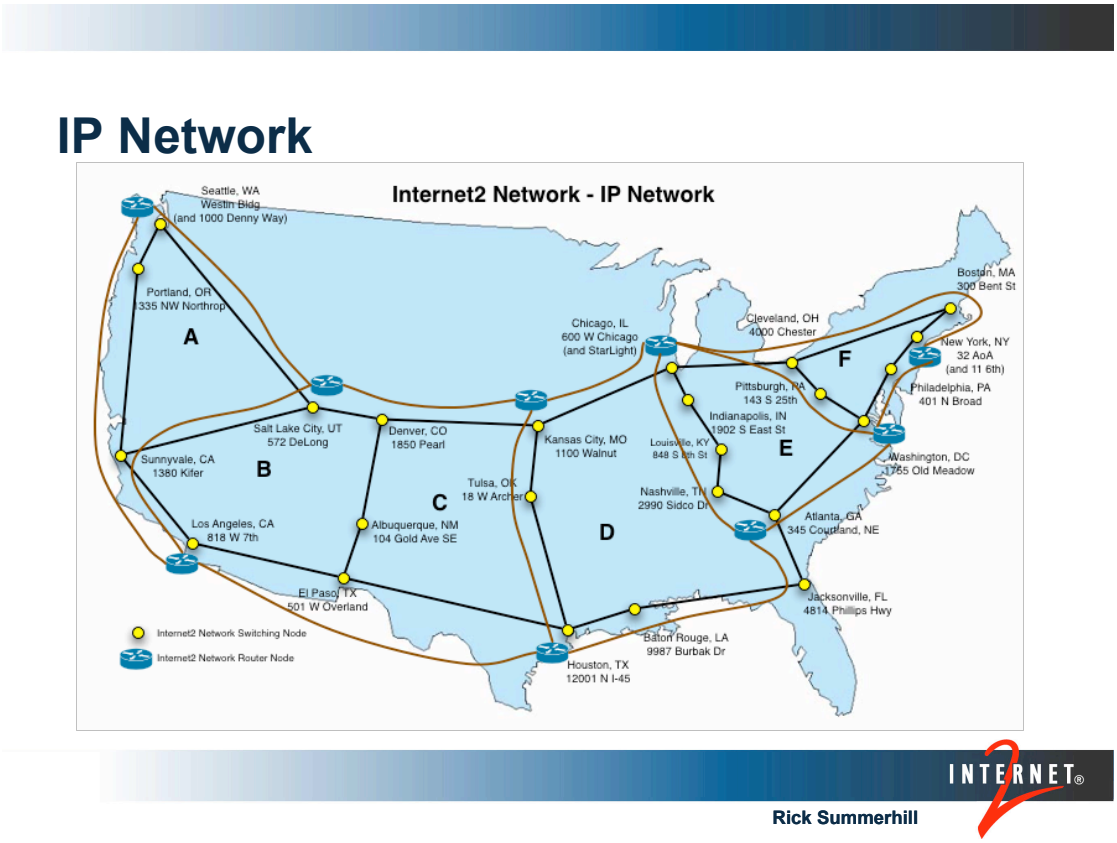

図 6.12: Internet2 IP Network（出典：Rick Summerhill, The New Internet2 Network, 2006.）

メインごとの計算を直列で行なっているが、並列計算を実行することで計算量の短縮が期待できる。

6.3 サブドメイン分割による SID テーブルエントリ削減

本節では階層型セグメントルーティング適用時のサブドメイン分割による SID テーブルエントリ削減効果の定式化を行う。

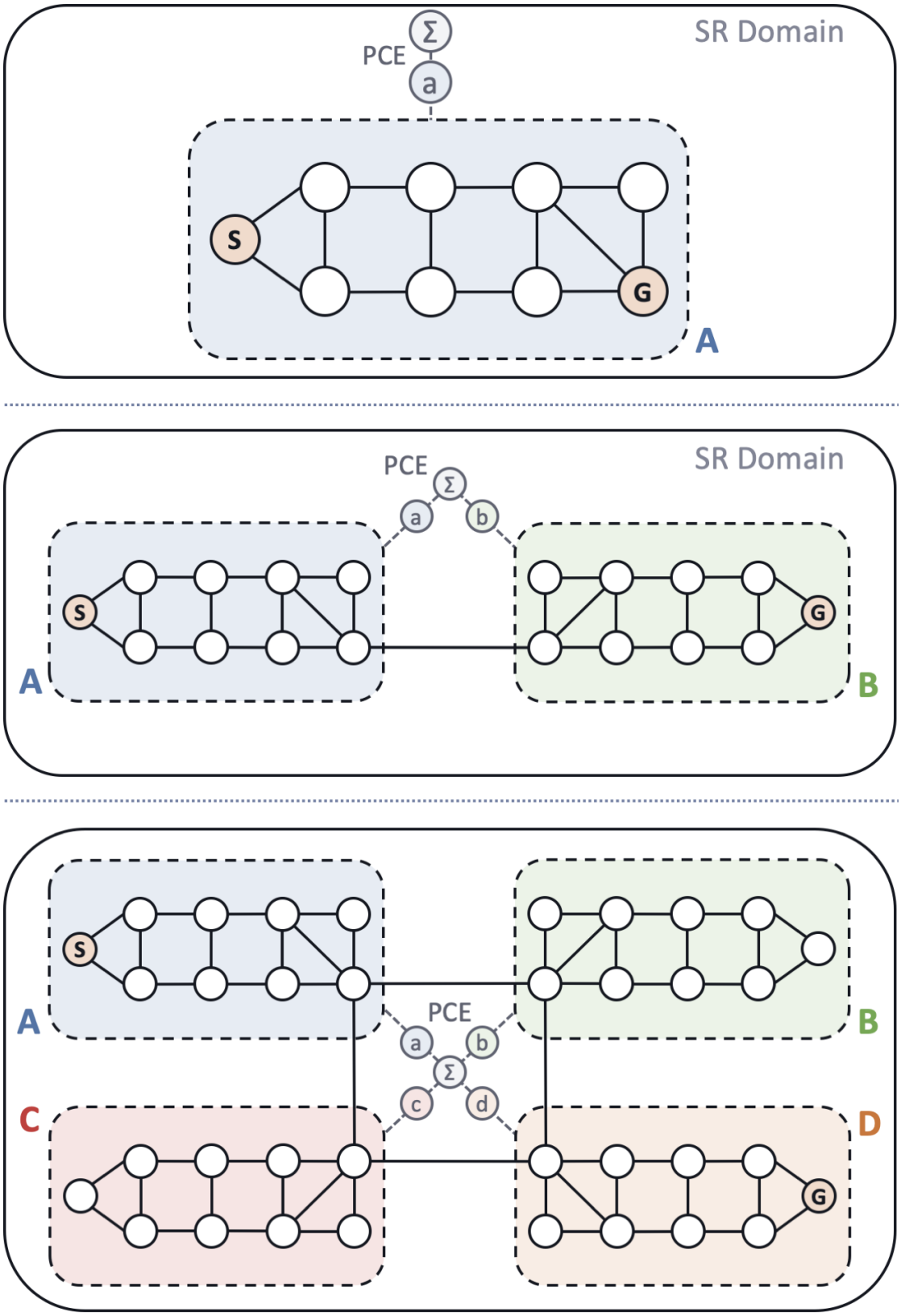

図 6.13: 計測用トポロジ

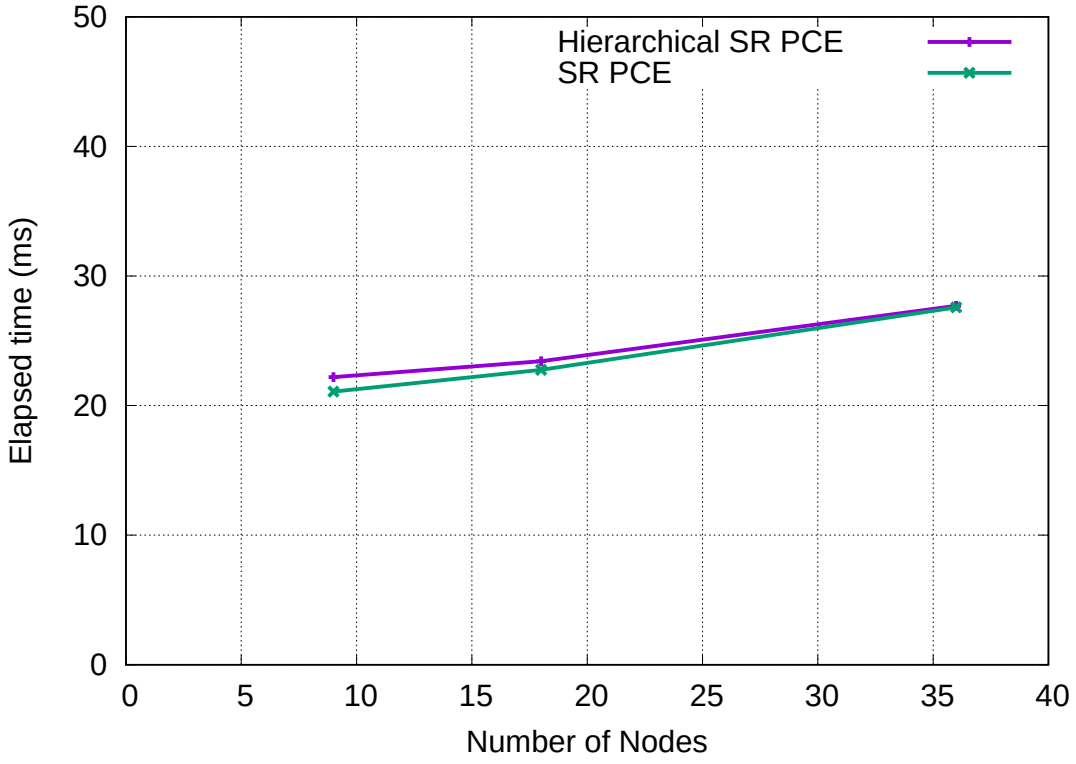

図 6.14: ノード数とセグメントリスト構築時間の関係

$$
\begin{aligned}
n &= All\ nodes \\
d &= Number\ of\ SR\ subdomain \\
s &= Node\ included\ in\ the\ subdomain \\
e &= Adjacency\ SID\ entry \\
t &= s + e - 1
\end{aligned}
$$

6.3.1 サブドメインの均等分割

第 4.1.1 項で扱った通り、サブドメイン分割の単位としては、AS ごとの分割や AS 内等のネットワーク単位での分割が考えられる。ここでは 1 つの AS を内部で分割する例を想定し、同規模のサブドメインに分割した際のネットワーク安定性とテーブルエントリ量について検討を行う。図 6.15 に 1 つの SR ドメインを 3 つのサブドメインに分割した例を示す。

図 6.15 の通り、SR ドメインには 12 個のノードが所属している。サブドメインサイズと SID テーブルエントリの相関を図 6.16 に示す。図 6.15 では、12 個のノー

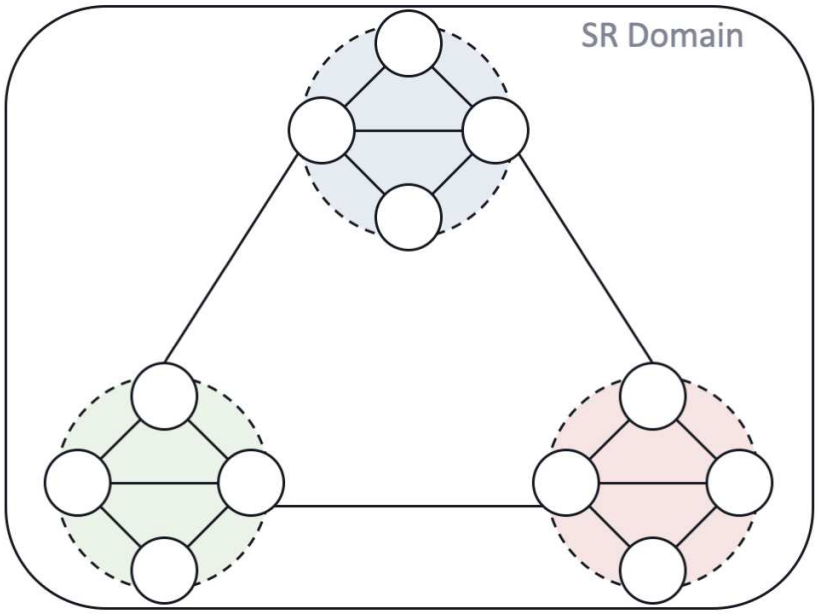

図 6.15: サブドメインの均等分割

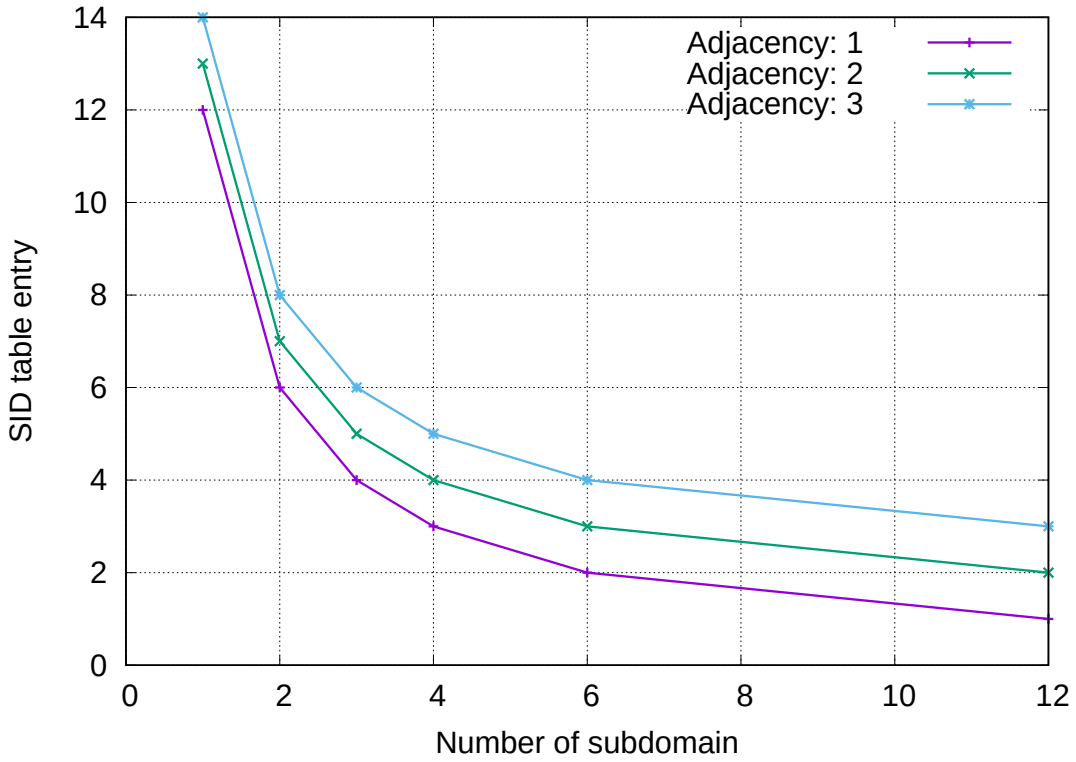

図 6.16: 均等分割でのサブドメインサイズとテーブルエントリの関係（n=10）

ドを 2・3・4・6・12 個のサブドメインに分割した際のそれぞれの SID テーブルエントリを示している。また、それぞれのグラフは各ノードの隣接関係がそれぞれ 1・2・3 本であったときのエントリ量を示している。図の通り、均等分割を行なった際サブドメインの個数が増加すると各ノードの SID テーブルエントリサイズが減少する。そのため、ネットワーク安定性の向上や規模追従性の向上を目指し同一のネットワークを分割する場合は、サブドメインを細かく分割した方が IGP の経路広告範囲やテーブルエントリサイズを削減でき、より安定性・規模追従性を向上できる。しかし、サブドメイン越えを想定した場合、いくつかの手法ではサブドメインを越えるたびにセグメントリストの増大が生じるため、パケットのヘッダサイズとのトレードオフが生じる。また、再発行モデルのようにサブドメイン越えの際にセグメントリストの付け替えが生じる場合は付け替えによる遅延が発生するため、同様にトレードオフが生じる。

6.3.2 ISP ネットワークにおけるサブドメイン分割

本章では、セグメントルーティングの規模追従性の評価として、1 つの ISP を内部のネットワーク単位でサブドメイン分割し、階層型セグメントルーティングを適用した例を想定する。図 6.17 に、ISP をコアネットワーク・アクセスネットワーク・ホームネットワークにわけモデル化した例を示す。図右側にはドメイン数 d、ノード数 n、ドメイン内のノード数 s を示している。

コアネットワークは ISP のバックボーンとなるネットワークである。ここでは各都道府県に 2 つのコアルータが存在すると想定し、台数を 100 台とした。アクセスネットワークはであり、それぞれのコアルータがアクセスネットワークを有しているものとする。ここでは各コアルータへ 100 台のアクセスルータが繋がっていると想定し、サブドメイン数 100、サブドメイン内のルータ数 100 台とする。ホームネットワークはである。日本電信電話株式会社の 2018 年度第 3 四半期決済補足資料 [22] によると、2018 年 12 月の契約数はフレッツ・ADSL が 819,000 回線、フレッツ光の契約数が 20,457,000 回線となっている。そのため、ここでは 1,000 万台のホームルータが存在すると想定した。またそれぞれのホームルータが 1 つのサブドメインを形成するため、サブドメイン数も 1,000 万個とする。

表 6.6 に、サブドメイン分割によるテーブル削減の効果を示す。既存のセグメントルーティングではあらゆるノードが全ノードの情報を保持するため、コア・アクセス・ホームの全てのネットワークにおいて、10,010,100 台のノードの情報を保持しており、規模拡大によるネットワーク安定性の低下や SID テーブルエントリ数による規模追従性の課題が予想される。それに対し階層型セグメントルーティングを適用した場合では、各ノードはサブドメイン内のノードのみを管理対象とする。具体的には、コアネットワーク・アクセスネットワークでは 100 台、ホームネットワークでは 1 台が対象となる。例より、ISP ネットワークにおいて、サブドメイン分割による管理対象ノードの削減効果が発揮されていることが確認できる。

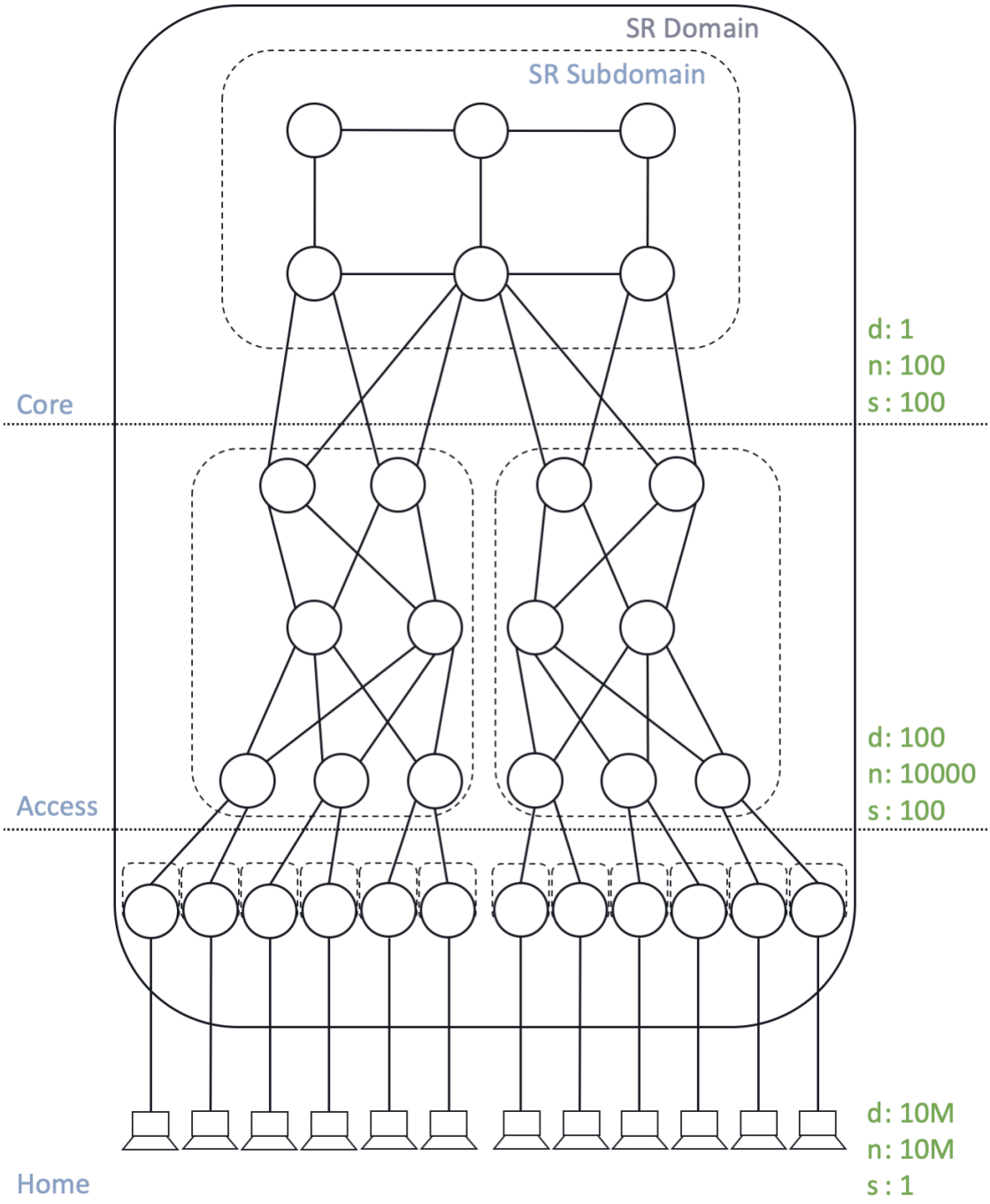

図 6.17: ISP ネットワークモデル

表 6.6: ISP ネットワークにおける SID テーブル削減の効果

Network Type	Table Entry (SR)	Table Entry (HSR)
Core	10010100	100
Access	10010100	100
Home	10010100	1

例より、階層型セグメントルーティングはノード数が大量であるほど SID テーブル削減が可能となり、安定性向上の効果を発揮することがわかる。しかし、収容ノード数が増加した場合、セグメントリスト構築を行う際の Dijkstra アルゴリズムの計算量も増加する。ISP ネットワークでは各ホームルータは他のホームルータの通信を中継しないため、Dijkstra アルゴリズムの計算対象はコアネットワーク・アクセスネットワークとなる。ここで 10,000 台規模までの ISP ネットワークにおけるセグメントリスト構築時間を予測する。

Dijkstra アルゴリズムの計算量は $\mathcal{O}(n \log n)$ であるため、近似式は $a * n \log n + b$ とし、nonlinear least-squares Marquardt-Levenberg algorithm を用いた。結果として、$a = 5.1354e - 05$、$b = 0.0209987$ が導かれた。

図 6.18 に予測結果を示す。結果より、10,000 ノードが存在する場合では 5 秒程度の経路計算時間が生じることが予測される。セグメントリスト構築時間の制約は SLA や用途などにより変化するため、制約を満たすように計算対象の削減を行う必要があると考えられる。これらの例により、セグメントルーティングの大規模利用が実現されたことを確認した。

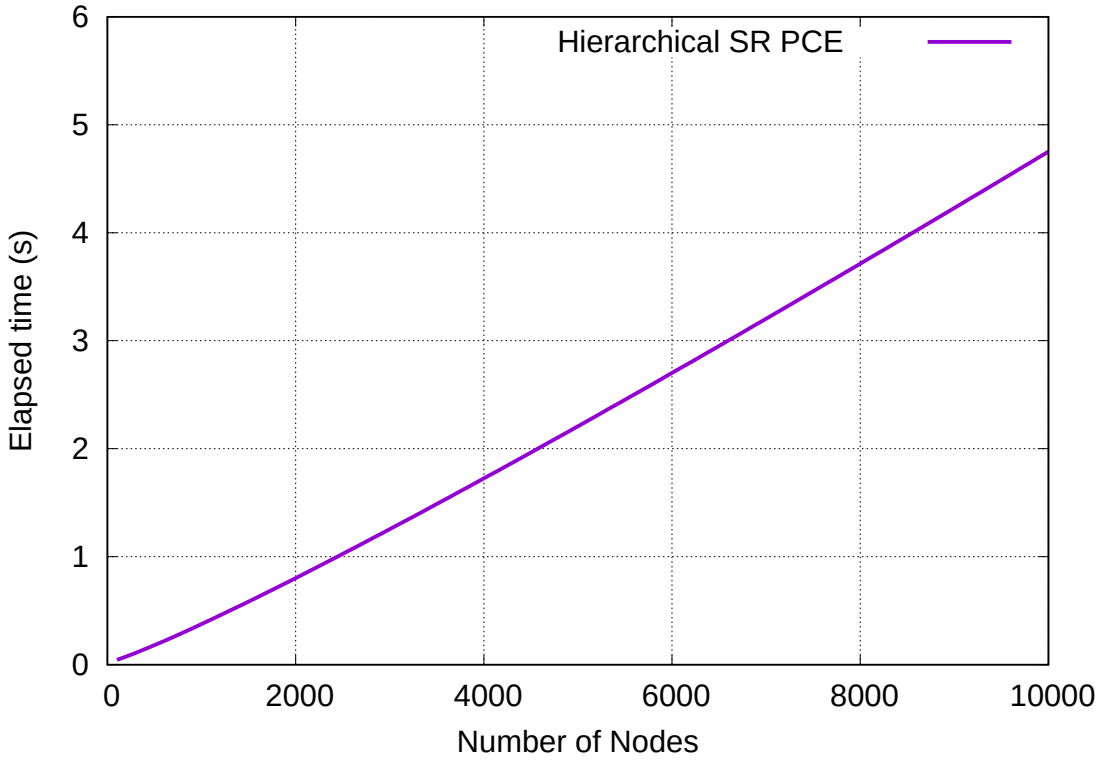

図 6.18: ISP ネットワークにおけるセグメントリスト構築時間の予測

第7章　おわりに

本章では、本研究における今後の課題・展望とまとめについて述べる。

7.1　今後の課題と展望

本節では、本研究に関する課題と今後の展望について述べる。

7.1.1　課題

本研究の提案手法である階層型セグメントルーティングにより、AS 連携環境へセグメントルーティングを適用する際に生じる管理の複雑性・規模追従性・安定性の課題を解決した。本手法に残る課題として、各 AS に存在するネットワーク機能の伝達の仕組みや、利用者が要求に応じてネットワーク機能を選択するための仕組みが必要となるだろう。

7.1.2　展望

本研究では対象としていないが、AS 連携におけるセグメントルーティングの改善策として、BSID の活用やデータプレーンの拡張によるセグメントリストの削減が考えられる。サブドメイン越えモデルのうち複数のモデルでは、トポロジやサブドメイン構成によりセグメントリストの増大が生じる。セグメントリストが増大した場合、パケットのヘッダサイズが増加するため、転送可能なデータサイズが減少し転送効率が低下する。セグメントリストの削減を行うことで効率的なデータ転送が可能となり、規模追従性の更なる向上が期待できる。

7.2　まとめ

本研究では複数の AS が存在するネットワークにセグメントルーティングを適用することを目指し、階層型セグメントルーティングの提案を行った。階層型セグメントルーティングでは、複数の AS が存在する環境へセグメントルーティングを適用する際に生じる管理の複雑化や規模追従性・安定性の低下を解決するため、

SR ドメインのサブドメイン分割を行い、階層化したコントロールプレーンを用いて経路計算を行う。本研究では経路計算の主体として PCE の利用を想定し、サブドメイン連携を可能とする階層型 SR PCE を提案した。また階層型 SR PCE の設計・実装を行い、AS 連携環境における性能評価を行った。

階層型セグメントルーティングの提案により、AS 連携環境におけるセグメントルーティング適用時の管理の複雑化と規模追従性・安定性の課題を解決した。これによりセグメントルーティングによる複数 AS の連携を容易にし、セグメントルーティングを利用した AS を越えるサービスチェイニングを実現した。

謝辞

本研究を進めるにあたり、多くの方に多大なご助言・ご助力をいただきました。それらの方々のご協力無くして、本研究は成り立ちませんでした。ここに深く感謝し、心よりお礼申し上げます。

指導教員である篠田陽一教授には多くのご助言とご指導を賜りました。深く感謝いたします。また、本研究室の知念賢一特任准教授、宇多仁助教には研究に関して活発な議論や多大なご指導を賜りました。深く感謝いたします。北陸 StarBED 技術センターの井上朋哉博士、明石邦夫博士には研究に関して様々な助言、ご指導を賜りました。心より感謝いたします。研究を進める上で助言をいただいた丹康雄教授、インターンシップ指導教員である Razvan Beuran 特任准教授に感謝いたします。

ソフトバンク株式会社 松嶋聡氏には研究を進める上でご助言と活発な議論をいただきました。ここに深く感謝いたします。また、WIDE Project の方々には研究を進める上でご助言と活発な議論をいただきました。深く感謝いたします。

本研究室修了生の可児友邦氏、三木晶司氏、押川侑樹氏、村上正樹氏には研究に関して活発な議論とご指導を賜り、また研究生活を送る上で様々なご助力をいただきました。心より感謝いたします。また、本研究室の博士後期課程の太田悟史氏、阿部博氏、三浦良介氏には、研究に関して活発な議論、ご指導を賜りました。心より感謝いたします。本研究室の博士前期課程の橋本光世氏、阿波史和氏、砂川真範氏、淺葉祥吾氏、広瀬太志氏、宮崎駿氏、小松源氏、山口礼央氏、菅野洋信氏、北沢尭宏氏、廣中颯氏、渡邊司揮氏には研究に関する活発な議論、研究生活を送る上での様々なご助力をいただきました。心より感謝いたします。

最後に、これまでの研究生活を支えてくれた家族、父 剛、母 薫、妹 愛と、祖父 隆、祖母 正子、伯父 毅、祖母 喜和子、叔母 綾子へ、そして天国の祖父 輝雄と叔父 哲へ心より感謝します。本当にありがとうございました。

本研究に関する対外発表

- 三島航, 明石邦夫, 宇多仁, 篠田陽一. セグメントルーティングにおけるセグメント階層化. マルチメディア, 分散協調とモバイルシンポジウム 2018 論文集, pp. 1397–1402. June, 2018
- 三島航. セグメントルーティングにおける階層化, WIDE Project ポスターセッション, September, 2018.
- 三島航. 01 より自由な経路制御を目指して, Sprouting Reserch –研究の芽–, `http://www.wide.ad.jp/Research/sr.html`, December, 2018.

参考文献

[1] Joel Halpern and Carlos Pignataro. Service function chaining (SFC) architecture. Technical report, 2015.

[2] Bo Han, Vijay Gopalakrishnan, Lusheng Ji, and Seungjoon Lee. Network function virtualization: Challenges and opportunities for innovations. *IEEE Communications Magazine*, 53(2):90–97, 2015.

[3] Clarence Filsfils, Nagendra Kumar Nainar, Carlos Pignataro, Juan Camilo Cardona, and Pierre Francois. The segment routing architecture. In *Global Communications Conference (GLOBECOM), 2015 IEEE*, pages 1–6. IEEE, 2015.

[4] Roger P Roess, Elena S Prassas, and William R McShane. *Traffic engineering*. Pearson/Prentice Hall, 2004.

[5] IETF Datatracker. Source packet routing in networking (spring). `https://datatracker.ietf.org/wg/spring/about/`.

[6] Bob Lantz, Brandon Heller, and Nick McKeown. A network in a laptop: rapid prototyping for software-defined networks. In *Proceedings of the 9th ACM SIGCOMM Workshop on Hot Topics in Networks*, page 19. ACM, 2010.

[7] Adrian Farrel and Jerry Vasseur, J-Pand Ash. A path computation element (PCE)-based architecture. Technical report, 2006.

[8] Siva Sivabalan, Jan Medved, Clarence Filsfils, V Lopez, J Tantsura, W Henderickx, E Crabbe, and J Hardwick. PCEP extensions for segment routing. *IETF draft-sivabalan-pce-segment-routing-03.txt*, 2014.

[9] Adrian Farral and Jerry Ash. A path computation element (PCE)-based architecture. *Request for Comments: 4655*, 2006.

[10] Manayya KB. Constrained shortest path first, August 2010.

[11] H. Gredler, Ed., J. Medved, S. Previdi, A. Farrel, and S. Ray. North-bound distribution of link-state and traffic engineering (TE) information using BGP, RFC7752. March 2016.

[12] Jean Philippe Vasseur and Jean Louis Le Roux. Path computation element (PCE) communication protocol (PCEP). Technical report, 2009.

[13] Ricard Vilalta, Ramon Casellas, Ricardo Martínez, Raul Muñoz, Young Lee, Haomian Zheng, Yi Lin, Victor López, and Luis Miguel Contreras. Fully automated peer service orchestration of cloud and network resources using ACTN and CSO. In *2018 International Conference on Optical Network Design and Modeling (ONDM)*, pages 124–129. IEEE, 2018.

[14] Clarence Filsfils, Siva Sivabalan, Daniel Yoyar, Mohan Nanduri, Steven Lin, Alex Bogdanov, Martin Horneffer, Francois Clad, Dirk Steinberg, Bruno Decraene, and Stephane Litkosky. Segment routing policy for traffic engineering. February 2017.

[15] Peter Psenak, Shraddha Hegde, Arkadiy Gulko, Clarence Filsfils, and Ketan Talaulikar. IGP flexible algorithm. 2018.

[16] Haijun Zhang, Na Liu, Xiaoli Chu, Keping Long, Abdol-Hamid Aghvami, and Victor CM Leung. Network slicing based 5G and future mobile networks: mobility, resource management, and challenges. *IEEE Communications Magazine*, 55(8):138–145, 2017.

[17] Clarence Filsfils. SR IGP flexible algorithm. `http://www.segment-routing.net/tutorials/2018-03-06-segment-routing-igp-flex-algo/`, March 2018.

[18] IETF Datatracker. Path computation element (pce). `https://datatracker.ietf.org/wg/pce/about/`.

[19] Dhruv Dhody, Young Lee, Lee Ceccarelli, Jongyoon Shin, Dan King, and Oscar Gonzalez de Dios. Hierarchical stateful path computation element (PCE). *IETF draft-dhodylee-pce-stateful-hpce-03.txt*, 2017.

[20] Dhruv Dhody, Young Lee, Lee Ceccarelli, Jongyoon Shin, Dan King, Oscar Gonzalez de Dios, and Xian Zhang. Stateful H-PCE&ACTN, March 2017.

[21] Rick Summerhill. The new internet2 network. In *6th GLIF Meeting*, 2006.

[22] 日本電信電話株式会社. 2018 年 3 月期 第 3 四半期 決算補足資料. `http://www.ntt.co.jp/news2018/1802jqqb/pdf/rbyv180209c_all.pdf`.